Reformulating active contours for object segmentation

Wang, Junyan

2012

Wang, J. (2012). Reformulating active contours for object segmentation. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/50709>

<https://doi.org/10.32657/10356/50709>

Reformulating Active Contours for Object Segmentation

Junyan Wang

School of Electrical and Electronic Engineering

A thesis submitted to the Nanyang Technological University
in partial fulfillment of the requirement for the degree of
Doctor of Philosophy

2012

To *my parents* and *my wife*

Acknowledgments

First and foremost I would like to express my sincerest gratitude to my supervisor, Dr. Kap Luk Chan, who has supported me throughout my PhD study with his patience and knowledge and whilst encouraging me to explore my own way of research. During these years of studying under his supervision, I have learnt a lot. I have been greatly impressed by his professional attitude toward research and teaching.

My PhD research began by taking two short courses recommended by Dr. Chan. One is the "Variational and PDE Techniques in Image Science", lectured by Professor Stanley Osher and Professor Xue-cheng Tai. The other is the "Image Analysis" given by Professor Alfred M. Bruckstein. The marvelous ideas, fascinating techniques shown by these enthusiastic professors shape my desired way of research. I would like to thank them sincerely.

I would like to thank Xinqi Chu for consistently reminding me the very pleasure of doing research by talking to me on various topics in machine learning, computer vision, mathematics and physics etc. I'm also indebted and thankful to Mao Li, Nini Liu, Jiwen Lu, Tiandi Duan, Xiaoyan Zhang and Ye Luo among many other kind people who have helped me and taught me a lot.

Finally, I would like to thank my parents who have persistently supported me during all my life. Special thanks must be given to my wife, Nan Wang, who has accompanied me during all these tough years. Without their support this thesis would not be possible.

Summary

Object segmentation is a fundamental problem in computer vision. Active contour is a general mathematical model for object segmentation. In an active contour model, an energy functional is formulated based on the observations of visual properties of object boundaries or object regions enclosed by the boundaries. By minimizing the energy functional, we expect to locate the enclosing boundaries of the regions of objects. This thesis addresses the following three fundamental problems in object segmentation based on active contour.

- (I) How can we formulate active contour models for segmenting objects based on its boundary properties or its regional properties?
- (II) If the object boundary or the object region appears partially in the image, due to occlusion or cluttered scene, how can we reformulate the active contour models in order to segment the object and recover the original object shape?
- (III) Given a real image containing a clutter of objects, how can we incorporate the prior information of the object of interest in the active contour formulation for segmenting the object of interest correctly without user intervention?

Most of the existing active contour models are based on edge detection, known as the edge-based active contours. The first contribution in this thesis is to address a well-known problem in the edge-based active contours, namely the problem of early termination of curve evolution. In this thesis, the problem is studied by examining the fundamental formulation of the edge-based active contour models. The formulation of edge-based active contour models often involves an edge indicator function. Conventionally, only the normal component of the gradient of the edge indicator function is considered. It is argued in the thesis that the tangential component should also be included. To include the tangential component, a new edge-based active contour model called the Geodesic Snakes and the corresponding curve evolution algorithm are developed for segmentation in edge-based active contour.

The active contour models based on regional properties have been proposed as an alternative to the edge-based active contour models. The second contribution in

this thesis is a novel region-based active contour model for coping with non-smooth image structures. For segmentation of non-smooth structures, such as textures, the image texture features have been previously used in the active contours. In this thesis, the problem of segmenting non-smooth image structures is formulated as a novel active contour model in terms of piecewise linear patch reconstruction, motivated by the Mumford-Shah model. It is also proven that the minimization of the patch reconstruction error leads to reduction of the segmentation error. Without requiring a separate process of feature extraction, the segmentation is accomplished in a principled way by minimizing the proposed single energy functional of piecewise linear patch reconstruction through curve evolution and optimal linear patch reconstruction.

Due to occlusion and/or cluttered scene resulting in missing parts or overlapping objects in the image, the energy minimization of the generic active contour models by curve evolution cannot extract the object shape of interest correctly. An object of interest may also undergo some deformation causing changes to its observed shape. Depending on the type of the object the deformation may be rigid or non-rigid. The third contribution of this thesis is a novel deterministic model of rigid and nonrigid shape deformations based on curve evolution, giving rise to the Prior Variation Shape Evolution (PVSE) equation. The PVSE is used in minimizing active contour energy for locating the object boundary of the deformed and/or recovered shape.

To segment the object of interest, an active contour is often initialized manually by an user. The fourth contribution in this thesis is a method to enable automatic contour initialization in active contours for segmenting the object of interest. By incorporating local feature matching and shape prior modeling into the active contour framework, a matching-constrained active contour model for automatic object segmentation is proposed. The initial contour, which is also the initial feasible solution to the constrained optimization, is generated by local feature matching followed by an affine invariant point-to-shape alignment.

The first and second contributions address Problem (I), the third contribution addresses Problem (II) and the last contribution tackles Problem (III). These contributions are made to the following three emerging directions of research on active contours: (i) the boundary and region modeling, (ii) the incorporation of shape priors and, (iii) the automation of object segmentation.

Contents

Acknowledgments	i
Summary	ii
List of Figures	ix
List of Tables	xix
List of Abbreviations	xx
List of Notations	xxi
1 Introduction	1
1.1 Motivation	1
1.2 Objectives	3
1.3 Major contributions of this thesis	4
1.4 Organization of the thesis	6
2 The Active Contour Framework	8
2.1 A brief literature review of active contours	9
2.2 The prototypical active contour models	11

2.2.1	Edge-based active contour models	11
2.2.2	Region-based active contour models	12
2.3	A unified view of edge-based and region-based active contour models	15
3	Active Contours with a Tangential Component	18
3.1	Introduction	19
3.2	Previous works on the early termination problem in curve evolution	23
3.3	The Geodesic Snakes model	25
3.3.1	Problem statement	25
3.3.2	The GeoSnakes and the Pseudo Stationary Phenomenon . .	26
3.3.3	The Equilibrium Flow and the Alternating Curve Evolutions	29
3.3.4	The geometrical analysis of the convergence of curve evolutions	33
3.3.5	The curve evolution algorithm for GeoSnakes	36
3.4	Experimental results	37
3.4.1	Experimental settings	37
3.4.2	Segmentation results	38
3.4.3	The tangential and normal velocities	43
3.5	Concluding remarks	44
4	Region-based Active Contour for Image Segmentation and Region Description for Non-smooth Image Structures	47
4.1	Introduction	48
4.2	Region-based active contours for unsupervised segmentation of non- smooth image structures	51
4.3	The principle of segmentation behind the Mumford-Shah model . .	52

4.4	Piecewise linear patch reconstruction	53
4.4.1	The model of piecewise linear patch reconstruction	54
4.4.2	The error of segmentation by piecewise linear patch reconstruction	55
4.4.3	Global optimal linear patch reconstruction	57
4.4.4	Coupling with piecewise smooth model	60
4.5	Experiments	61
4.5.1	Evaluation of the gradient descent patch reconstruction	62
4.5.2	Evaluation of the segmentation by linear patch reconstruction	63
4.5.3	Segmentation by the coupled model with parameter tuning	65
4.5.4	One-against-all segmentation	69
4.6	Discussions	70
4.7	Concluding remarks	71

5	Rigid and Nonrigid Shape Evolutions for Shape Registration and Recovery	72
5.1	Introduction	73
5.1.1	Related works	74
5.1.2	Contributions	76
5.1.3	Organizations	78
5.2	Modeling shape variations by curve evolutions	79
5.2.1	Infinitesimal Warping Iteration	79
5.2.2	Prior variation shape evolution	81
5.3	A theory of shape preservability	83

5.3.1	The invariance of shape characteristics	84
5.3.2	The prior vibrations	86
5.4	Calculus of prior variations	87
5.5	Experimental results	90
5.5.1	Evaluation of the PVSE equation for modeling shape deforma- tions	90
5.5.2	Evaluation of the PVSE for shape registration	95
5.6	Discussions on the limitations	98
5.7	Concluding remarks	99
6	Matching-Constrained Active Contours	101
6.1	Introduction	102
6.1.1	Related works	102
6.1.2	Methodology	104
6.1.3	Contributions	107
6.1.4	Organization	108
6.2	Modeling affine-invariant interior-points-to-shape relation	109
6.2.1	Points-to-shape relation as a binary classifier	109
6.2.2	Imposing affine invariance in the points-to-shape relation . .	111
6.3	The matching-constrained active contour	115
6.3.1	Active contour model with matching constraint	116
6.3.2	Projected gradient decent algorithm	117
6.3.3	Gradient of active contour parameterized by affine motion of interior points	118

6.3.4	Joint formulation of point matching and alignment	119
6.4	Experiments	120
6.4.1	Affine invariant shape modeling	121
6.4.2	The importance of good initialization to active contours . . .	124
6.4.3	The fundamental problem of global search	125
6.4.4	Automatic object segmentation by matching-constrained ac- tive contour	128
6.4.5	Discussions	130
6.5	Concluding remarks	132
7	Conclusion and Recommendations	134
7.1	Conclusion	134
7.2	Recommendations for future research	136
	Appendices	138
A	Gradient descent method for active contours	138
A.1	Functional gradient and gradient descent	138
A.2	The calculus of variations for active contours	143
B	Appendix of Chapter 3	153
C	Appendix of Chapter 4	162
D	Appendix of Chapter 5	171
E	Appendix of Chapter 6	175
	Author's Publication	178
	Bibliography	179

List of Figures

1.1	Illustration of the problems.	3
3.1	The full gradients of the edge indicator near the object boundary are close to zero. (a) is an image containing an object of interest. In (b), the center is the edge indicator function for (a). The intensity of the edge indicator is visualized by the darkness in figure (b). The surrounding patches in (b) show the gradient vector field of the edge indicator along the object boundary (dashed curve).	21
3.2	The U-shape and the opposite gradient vectors	24
3.3	The curve stops (or oscillates) when the gradient is nearly orthogonal to the inward normal of the curve. (a) is a 3D visualization of a typical edge indicator function overlayed by the negative gradient field; (b) demonstrates the curve evolution in the negative gradient field within the gray-shaded box region in (a).	26

3.4	The practical effect of rotating the gradient field. The top left is a typical gradient field, visualized by ball blue arrows, in which the green line segment cannot be moved by the local gradients due to the PSP. The top right is the vector field generated by the rotation, visualized by the arrows in pinkish red, in which the line segment can now be moved by this vector field. The bottom visualizes both the gradient field and the rotated vector field. It shows that the stationary positions, i.e., the light blue curve, remain stationary after the rotation	32
3.5	Segmentation of the U-shape. (a) is the initialization. (b) is the PSP. (f) is the converged result.	38
3.6	Inputs and initializations	39
3.7	The original real images and the tumor regions (in boxes)	40
3.8	Gradient fields, PSP and Saddle points	40
3.9	Object Extraction by proposed method. The first row is the curve evolution by GAC; The second is the curve evolution by EF; The bottom row is the final convergence.	41
3.10	Results by other methods. First column from the left is the boundary extraction by GAC; Second column is the one by GAC with Balloon; Third column is the one by GAC with adaptive Balloon; Fourth column is the one by MAC; Last column is the one by Chan-Vese active contour.	42
3.11	The reconstructed 3D tumor from the MRI sequence of Figure 3.6(e).	43

3.12	The normal and tangential velocities during curve evolutions. The top row shows the total magnitude of the tangential velocity along the curves; The second row shows the total magnitude of the normal velocity along the curves. k_1, k_2, k_3, k_4 are some key iterations in the curve evolutions. 0- k_2 is the curve evolution by GAC; k_2 - k_4 is by EF; from k_4 onwards is by GAC	44
4.1	Illustration of the problem. The left two images are composed of different image structures that the conventional Mumford-Shah model would fail to segment; The diagram on the right shows that the Gabor feature image of a texture is non-smooth.	49
4.2	(a) is the template for texture mosaicing. (b) is the initial contour for curve evolutions.	62
4.3	The scatter plots of the <i>normalized</i> reconstruction errors for each face or texture image by the first 1-20 bases. The value 1 corresponds to the worst reconstruction, and 0 corresponds to the perfect reconstruction. The left two plots are the orthogonalized Gabor bases (vertical) vs. SVD (horizontal); The other two are the Gradient descent (vertical) vs. SVD (horizontal). Results by gradient descent are almost as the same as SVD for different numbers of bases. . . .	64
4.4	Comparison of segmentation error rates on the images with the mean intensity subtracted (left), and on the images of original textures (right). Proposed methods, i.e. the SVD & GD, are robust and more accurate.	66

4.5	Curve evolution on the Brodatz D104-22 pair (top row) and D85-106 pair (mid row) and a real image taken from Berkeley dataset with the corresponding converged patch bases. The curves are drawn in red (better to view in color). The leftmost and rightmost two columns of the patch bases correspond to the background/foreground regions. The D22, D106 and the lady are regarded as the foregrounds.	66
4.6	The converged curve evolution results of different methods with the same initialization	67
4.7	The box plot of segmentation error against α	68
4.8	The segmentation of synthetic piecewise smooth image under noise. The images on the left show the initialization, the images in the middle show the converged curves. In the images on the right, the left two columns are the bases of the background region, and the rest correspond to the vessel region.	68
4.9	The segmentation of piecewise smooth images. The images on the left show the initialization, the images in the middle show the converged curves. In the image mosaics on the right, the left first two columns of mosaics from the left are the bases of the background region, and the rest correspond to the vessel region.	68

4.10	The segmentation of piecewise smooth images by purely linear patch reconstruction. The top row shows the results of segmentation. The bottom row shows the optimal bases corresponding to the images above. In the image mosaics of the bases, the left two columns are the bases of the background region, and the rest correspond to the foreground region (the regions enclosed by the contour curves). . . .	69
4.11	Multi-region segmentation via the one-against-all strategy (better to view in color). The top row shows the two input-output pairs of mosaic images composed of multiple texture regions. The input is the initial rectangular contour laying over the image. In each region, a unique color is assigned to the contour. The output is the converged curve shown in the same color as in the input. The bottom rows show the optimal bases corresponding to each region in the sequence of top, bottom, left, right and center.	70
5.1	(a) A toy monkey; (b) toy monkey after a rigid transformation; (c) toy monkey after a nonrigid transformation; (d), (e) and (f) are the toy monkeys in (a) (b) and (c) with missing parts.	73
5.2	(a) A horse; (b) The rigidly transformed horse; (c) the nonrigidly deformed horse; (d), (e) and (f) are the horses in (a), (b) and (c) merged with the rider.	74
5.3	A horse registration process	74
5.4	The curve evolution for concurrently rotation, scaling and translation of “A”.	83
5.5	The curve evolution for affine transformation of “A”.	83

5.6	The prototype (left) and baseline (middle) shapes used for evaluation, and the right shows the affine shape context matching of the baseline (blue crosses) by the prototype (red circles).	92
5.7	The shape distance for the 50 iterations of similarity (left) and affine (right) PVSEs. The red line is the shape distance between the two “A”s in Figure 5.6.	93
5.8	The shape context distance for the 50 iterations of the Prior Vibration Curve Evolutions of order 1 to order 10. The left shows the shape distances for each order of prior vibration; The right shows the shape distances at each iteration for all orders. The red line is the shape distance between the two “A”s in Figure 5.6.	94
5.9	The PVSE and the affine shape matching by shape context. The blue curves are the contour curves due to PVSE, the red crosses are the results of shape matching by shape context. The first row corresponds to a PVSE by similarity variations; The second row corresponds to a PVSE by affine variations; The last row corresponds to a PVSE by 3 rd order prior vibrations.	95
5.10	Shape Variation: the distances between the initial prototype shapes and the shapes to be registered. This is to show that the deformations of shapes is relatively large. The task is challenging	97
5.11	Registration performance measured by Jaccard coefficients (vertical axis).	98
5.12	PVSE with prior vibrations for registration of the toy monkey. (a) to (e) show the curve evolution of 90 iterations until convergence at (e).	98

5.13	PVSE with prior vibrations for registration of the horse shape. (a) to (e) show the curve evolution of 150 iterations until convergence at (e).	98
5.14	PVSE with prior vibrations for registration of another horse shape. (a) to (e) show the curve evolution of 500 iterations until convergence at (e).	99
5.15	Shape registration with affine PVSE to find the “ $\hbar$ ” in the tattoo on the skin. (a) to (e) show the curve evolution of 200 iterations until convergence at (e).	99
5.16	Segmentation and recovery of the TPAMI logo from heavy noise and occlusion by PVSE with the 6 th -order prior vibration. The top-left is the input image. The bottom-left is the convergent curve laying over the ground truth. (a) to (h) visualize the curve evolution of 150 iterations until convergence at (h).	99
5.17	The box of the clownfish overlaying the original color image and the red channel to which the PVSE applies.	100
5.18	Shape registration with PVSE to find the clownfish in the coral. The first row shows the affine shape evolution until convergence. The second row shows the prior vibration shape evolution until convergence.	100
6.1	The importance of initialization. (a) shows a bad initialization which gives an unsatisfactory result of the segmentation in (b). (c) shows a good initialization which gives an encouraging result in (d).	105
6.2	Illustration of the feature point matching (with errors).	106

6.3	Object segmentation through active contour with shape prior initialized by feature point matching.	106
6.4	A 2D visualization of the trained shape decision function overlayed the corresponding implicit contour (in red). The larger the intensity of the image is the more likely the point belongs to the shape region.	111
6.5	Illustration of the need of affine invariance. The point correspondences should lead to the shape alignment.	112
6.6	Significance of affine invariance of the shape representation. Blue dots are the transformed feature points. The outlining curves are the implicit shape contours.	114
6.7	A demo of points-to-shape alignment.	115
6.8	Three pairs of template objects (left) and the corresponding trained shape contours overlaying the training shapes (right)	122
6.9	Minimization of the errors of shape modeling corresponding to Figures 6.8 (a) (b) and (c).	122
6.10	Experimental results on selection of the parameter σ . (a) is the plot of fitting scores (vertical) v.s. σ (horizontal), given the fixed initial weights and thresholds. (b) is the plot of fitting scores w.r.t. different values of σ given the fixed initial weights and thresholds (blue square with curve) comparing with the fitting scores due to the optimization of the weights and threshold given the randomly selected σ (red cross with curve).	123

6.11	The significance of the affine invariance of the interior-points-to-shape relation. The top row shows the shapes generated by the affine-invariant model in Eq. (6.8). The bottom row shows the shapes generated by the model without affine invariance in Eq. (6.7). The blue dots are the interior feature points, the red curves are the implicit contours, and the black circles are the explicitly transformed contours.	125
6.12	The normalized Hausdorff distances between the shapes implicitly defined by the transformed interior points and the original shapes (horizontal) v.s. the distance between the explicitly transformed contour shapes and the original shape (vertical).	126
6.13	The object segmentation by the affine active contour driven by affine motion of interior points. (a) is a bad initialization which leads to a bad result shown in (b). (c) displays a good initialization which leads to a good result shown in (d).	127
6.14	The object segmentation by exhaustive search in simple image. (a) is the result by Chan-Vese model (b) is the result by GAC model .	127
6.15	The object segmentation by exhaustive search in a cluttered image. (a) shows the result by the Chan-Vese model (b) shows the result by the GAC model	128
6.16	Segmentation of an object of different poses in noisy images in the presence of occlusion by leaves, clutter background and unclear boundary on the top. The rows of the figures from top to bottom are referred to as (a) (b) (c) and (d).	130

6.17	Segmentation of an object of relatively complex shape under different poses and scales with non-ideal initial matchings. The top and bottom triple of figures are labeled as (a) and (b).	131
6.18	Segmentation of an object of complex shape under different scale, clutter background and indefinite boundary.	131
6.19	Quality of segmentation (vertical) v.s. score of model fitting (horizontal)	132
A.1	Contour splitting.	140
A.2	The idea of the level set method. The dotted curves on ϕ are the zero level sets	141

List of Tables

3.1	Timings for curve evolutions on images of different sizes	45
4.1	Comparison of segmentation error summarized from Figure 4.4. . .	65
4.2	Errors of the segmentation shown in Figure 4.11	70
5.1	Haussdorff shape distance corresponding to Figure 5.9	94
5.2	The distributions of randomly generated parameters from uniform distributions	96
5.3	Abbreviations	97
6.1	Shape distances corresponding to the Figure 6.11.	124
6.2	The Jaccard similarity against the ground truth.	129
6.3	Computational costs (in seconds)	129

List of Abbreviations

EF	Equilibrium flow
EL	Euler-Lagrange
GAC	Geodesic active contour
GD	Gradient descent
GVF	Gradient vector flow
IWI	Infinitesimal warping iterations
LBF	Local binary fitting
MAC	Magnetostatic active contour
MRI	Magnetic resonance imaging
NGVF	Normalized gradient vector flow
PC	Piecewise constant
PCA	Principal Component Analysis
PS	Piecewise smooth
PSP	Pseudo stationary phenomenon
PVSE	Prior variation shape evolution
RBF	Radial basis function
SIFT	Scale-invariant feature transform
SVD	Singular value decomposition

List of Notations

$C(p)$	A curve parameterized by p
$C(p, t)$	Evolving curve C parameterized by p and artificial time t
$\vec{T}(C(p, t))$	Tangent of curve C at position p and time t
$\vec{N}(C(p, t))$	Normal of curve C at position p and time t
$\partial\Omega$	Boundary of region Ω
κ	The mean curvature of a curve
$\ \vec{v}\ $	The norm of vector $\vec{v}$
$\langle \vec{a}, \vec{b} \rangle$	Vectorial inner product of vectors $\vec{a}$ and $\vec{b}$
$\ f\ _{L_p}$	The L_p -norm of function f
$\langle f, g \rangle_{L_p}$	L_p inner product of functions f and g
g_p	The derivative of function g with respect to p
∇g	The gradient field of scalar field g
$\frac{\partial g}{\partial p}$	The partial derivative of function g w.r.t. p
$\frac{dg}{dp}$	The total derivative of function g w.r.t. p
$\nabla \cdot \vec{v}$ or $\text{div}(\vec{v})$	The divergence operator of the vector field of $\vec{v}$
$\alpha \in A$	α is an element of set A
$\alpha \notin A$	α is not an element of set A

Chapter 1

Introduction

Computer vision research tries to answer the fundamental question: how might computers see? This area is flourishing, but the computer vision community has yet to reach some consensus regarding various issues in emulating biological vision by computational means. While computational vision not necessarily exactly follows the biological system, a lack of consensus encourages diverse approaches to the vision problems. Theorists approach the vision problems by exploring the modeling of them while engineers tend to approach them from the viewpoint of the application in which such problems exist and develop both hardware and software solutions accordingly.

1.1 Motivation

One of the open problems in computer vision is the extraction of regions of objects from an image, i.e., object segmentation. For example, for any given natural image, human can extract regions of objects and each individual region contains precisely a single object. The question is whether a machine or an algorithm can achieve the same segmentation. This problem can be straightforwardly posed as a Turing test of vision machines.

Currently, object segmentation becomes a prevailing research topic in computer vision. Object segmentation has been formulated as energy minimization problems, among which the active contour model is an influential one. The first active contour model of object segmentation was developed systematically based on the early vision process of edge detection. Early vision processes recover physical properties of visual objects, and they were often believed to be independent of domain-specific knowledge [1]. Poggio et al. [2] developed a theory of computational vision by posing a variety of early vision tasks as mathematical ill-posed problems which were solved by using the regularization theory. Differing from the original view of the early vision, this theory of computational vision incorporated *a priori* knowledge represented by the regularization. Originally, Poggio did not formulate the problem of object segmentation to follow the regularization theory of computational vision. This was done by Kass, Witkin and Terzopoulos in [3], in which they integrated Marr's theory of edge detection [4] with the regularization theory of computer vision to establish an energy minimization model of object segmentation. By minimizing the energy, the object boundary, which is a 2D curve, can be computed. This method by Kass et al. was named the "Snakes" or "active contour".

In spite of the sound regularization principle in active contours, we wonder why active contours are well-suited for object segmentation. One of the advantages is that the active contour model is a mathematical model of object segmentation, contrasting many heuristic computational frameworks. The quality of the segmentation is measured by active contour energy, which establishes an objective formulation of the problem of object segmentation. This energy measure allows the assessment of the solution without the subjective evaluation by humans, which is the ultimate goal of object segmentation by computer. Another advantage is that the formulation of the active contour energy is flexible, and the numerical solution of the energy minimization is well-established, which allows users to de-

Figure 1.1: Illustration of the problems.

velop their active contour models and algorithms to suit the characteristics of their images. There exists many applications of active contours in literature, such as in medical image analysis [5–7], surveillance video analysis [8–10].

1.2 Objectives

It is still not very certain to what extent that active contours can be useful. Hence, to explore further possibilities of object segmentation with active contours, the limitations of the current framework of active contour are to be identified and addressed. This thesis is devoted to some fundamental problems encountered with the active contour framework in order to improve the methodology and to extend the original scope of active contours for object segmentation.

This thesis addresses the following three fundamental problems of the active contour framework.

- (I) How can we formulate active contour models for segmenting objects based on its boundary properties or its regional properties? For instance, what is the proper model for the tumor segmentation in Figure 1.1(a) and for differentiating the two different non-smooth structures in Figure 1.1(b)?
- (II) If the object boundary or the object region appears partially in the image, due to occlusion or cluttered scene, how can we reformulate the active contour models in order to segment the object and recover the original object

shape? For instance, how could we extract the shape of the clownfish in Figure 1.1(c)?

- (III) Given a real image containing a clutter of objects, how can we incorporate the prior information of the object of interest in the active contour formulation for segmenting the object of interest correctly without user intervention? For instance, how could we segment out the toy guard from the cluttered scene in Figure 1.1(d) automatically?

The three problems are interconnected. Problem (I) concerns with the fundamental rationale of the active contour framework. It is implicitly assumed that there exists such a measure for locating the object boundary or region. Problem (II) actually asks how to estimate the boundary or region of the object when the image data does not fully agree with the formulated active contour model, due to erroneous model or missing data. Besides, there could be many objects in an image and some of them might not be of interest. Problem (III) asks how to identify the object of interest without user intervention in active contours in order to achieve automatic object segmentation, assuming that the object model is available and roughly correct. The Problems (II) and (III) could be viewed as the natural outcome of the evolution of the research initiated by addressing Problem (I).

1.3 Major contributions of this thesis

Most of the existing active contour models are based on edge detection, which are now known as the edge-based active contours. The first contribution in this thesis is to address a well-known problem with the edge-based active contours, namely the problem of early termination of the curve evolution. In this thesis, the problem is studied by examining the fundamental formulation of the edge-based

active contour models in which only the normal component of the gradient of the edge indicator function is considered. It is argued in the thesis that the tangential component should also be included. To include the tangential component, a new edge-based active contour model called the Geodesic Snakes (GeoSnakes) and the corresponding curve evolution algorithm are developed for segmentation.

The active contour models based on regional properties, a.k.a. the region-based active contour models, have been proposed as an alternative to the edge-based active contour models. The second contribution in this thesis is a novel region-based active contour model for coping with non-smooth image structures. Previously, the region-based active contours have been applied to the feature image for segmentation of non-smooth structures, which is however heuristic. In this thesis, the problem of segmenting non-smooth image structures is formulated as a unified mathematical model in the form of piecewise linear patch reconstruction, motivated by the Mumford-Shah model. The segmentation is attainable by minimizing the single proposed functional energy without any separate process of feature extraction. To avoid heuristics, it is also proven that the minimization of the patch reconstruction error leads up to reduction of the error of segmentation.

Due to occlusion, cluttered background, resulting in missing parts or overlapping shapes in the image, the energy minimization of the generic active contour models by curve evolution cannot extract the underlying object shape. The third contribution of this thesis is a novel deterministic deformable shape model based on curve evolution, giving rise to the Prior Variation Shape Evolution (PVSE) equation. The PVSE can be used to minimize active contour energy for location of the object boundary and for recovery of the underlying object shape.

To segment the object of interest, initializing an active contour by an user is often required. The fourth contribution in this thesis is to enable automatic contour initialization in active contours. By incorporating local feature matching and

shape prior modeling into the active contour framework, a matching-constrained active contour model for automatic object segmentation is proposed. The initial contour, which is also the initial feasible solution to the constrained optimization, is generated by local feature matching followed by an affine invariant points-to-shape alignment.

The first and second contributions addresses the Problem (I), the third contribution addresses Problem (II) and the last contribution addresses Problem (III).

1.4 Organization of the thesis

The remaining chapters of this thesis are organized as follows. In Chapter 2, a brief literature review of the active contour and its mathematical background related to this thesis are presented. More detailed mathematical treatment regarding the energy minimization for active contours is deferred to the Appendices. The subsequent four chapters are the main contributions of this thesis.

In Chapter 3, A new edge-based active contour model called the GeoSnakes is formulated, and the numerical solution to the GeoSnakes is based on a new type of curve evolution formulation named the *equilibrium flow*. Chapter 4 addresses the segmentation of non-smooth image structures by active contours based on regional information. The segmentation is accomplished in a principled way by minimizing the proposed single energy functional of piecewise linear patch reconstruction through curve evolution and optimal linear patch reconstruction. Chapter 5 addresses the shape extraction by active contour under occlusion, cluttered background and/or low contrast, which often result in missing parts or overlapping objects in the image. A novel deterministic model of rigid and nonrigid shape deformations based on curve evolution is proposed, giving rise to the Prior Variation Shape Evolution (PVSE) equation. The PVSE is used to minimize active contour

energy for locating the object boundary of the deformed and/or recovered shape. In Chapter 6, the automatic object segmentation is addressed. By incorporating local feature matching and shape prior modeling into the active contour framework, a novel active contour model for automatic object segmentation is proposed. The matching based initialization and the curve evolution is unified through a novel formulation of constrained optimization, which is named the matching-constrained active contour.

Chapter 7 provides the conclusion on the thesis with discussions on the future work. The Appendices contain lengthy or less important proofs and derivations of the new mathematical claims and new equations developed in this thesis.

Chapter 2

The Active Contour Framework

In this chapter, the framework of active contours for object segmentation is introduced. An excellent introduction of curve evolution with some necessary differential geometry can be found in [11]. Some recent developments of level set method can also be found in [12]. A text book of the level set method covering the major numerical technique adopted in this thesis is [13].

For the problem of object segmentation in images, an object can be defined by the boundary enclosing the object. The idea of active contour was proposed based on this observation. The requirements of applying active contour in object segmentation is summarized below.

- (i) Find a mathematical representation of any closed contour curve.
- (ii) Based on the representation, formulate an optimization problem for which the boundaries of objects are expected to be given by the optimal solution.
- (iii) Apply appropriate optimization methods to the optimization problem.

The mathematical details on these issues can be found in the Appendix A.

2.1 A brief literature review of active contours

Kass et al. [3] proposed the first active contour, the Snakes, based on the regularization theory of computational vision [2] and Marr's theory of edge detection [4]. It replaces the traditional heuristic seed growing and region merge/split strategies with the energy-minimizing curve evolutions. However, the original method requires explicit reparametrization of the contours. Meanwhile, in the area of computational physics, Osher and Sethian [14] proposed an implicit level set method for simulating front propagation, which is numerically more advantageous than the conventional explicit parametric schemes adopted in the original Snakes. Caselles et al. [15] and Malladi with Sethian [16] tried separately to improve the numerical solution of the active contours by using the level set method. However, the level set method is not very compatible with the original formulations of the Snakes. They derived new curve evolution equations that can be solved by the level set method. Later, Caselles et al. [17] and Yezzi et al. [6] separately obtained the geometrical formulation of the active contours, which are fully compatible with the level set method.

The first region-based active contour, known as the region competition, was proposed by Zhu and Yuille [18], though the heuristic strategy of region-based curve evolution was explored earlier in [19]. Both of the works were motivated by [20] in which a general formulation of the idea of segmentation by piecewise reconstruction (termed the visual reconstruction by Blake and Zisserman in [21]) was obtained. The idea of segmentation based on piecewise smooth image reconstruction can be traced back to the works of Pavlidis [22] and Zucker [23] in which it was recognized that the region growing is equivalent to piecewise smooth function approximation. The level set method was first applied to solve a simplified Mumford-Shah model, i.e., the Mumford-Shah cartoon functional, by Chan and Vese in [24]. The resultant active contour is known as the Chan-Vese active con-

tour. More recently, researchers are interested in the linear relaxation [25] or the convex relaxation [26] of the full Mumford-Shah model to cope with the numerical difficulties in the minimization of the original energy functional. Another convex relaxation of the full Mumford-Shah functional was proposed in [27], in which the Mumford-Shah functional is considered as a pure reconstruction model. Brox and Cremers [28] attempted to interpret the full Mumford-Shah model as a Bayesian inference model. There are also other region-based variational formulations of segmentation, such as the continuous max-flow/min-cut formulations [29, 30].

In order to correctly outline the boundary of the object of interest in the presence of occlusion and cluttered background in the image, some higher level knowledge, such as a prior shape of the object, is needed. Chen et al. [31] proposed a variational curve evolution framework for object segmentation by aligning the evolving curve with a prior shape iteratively. In the context of shape registration, this work only considered the rigid deformations. Leventon et al. [32] modeled the nonrigid shape variation using Principal Component Analysis (PCA), and they measured the shape distance using Gaussian density centered at a mean shape. Cremers et al. [33, 34] developed a framework involving both variational rigid shape alignment and statistical models. More recently, Etyngier et al. [35] formulated the space of prior shapes for the curve evolution to be a shape manifold. This formulation generalizes the previous statistical prior shape models.

More recently, Schoenemann and Cremers [36] proposed a global optimization method for the active contour taking account of patch comparison, edge detection and translation-invariant shape prior. The method achieves automatic object segmentation with multiple reference shapes, under the assumption that the object shape is sufficient to specify the object of interest.

2.2 The prototypical active contour models

2.2.1 Edge-based active contour models

Informally, the idea of the Snakes was to find a 2D curve represented by a 2D vectorial function to capture the boundary edge of a targeted object. By modeling the boundary of the object as the optimal solution of an optimization problem, the boundary curve can then be found by means of numerical optimization. The algorithm is conventionally the gradient descent method upon given an initial curve. During the gradient descent, the curve is deforming before the optimal solution is reached. The formulation of the Snakes is as follows. Firstly, a closed curve was considered as a vectorial function of a single parametrization variable, i.e., $C = C(p) = [x(p), y(p)]^T$. Then a *functional energy minimization* in terms of $C(p)$ was formulated as follows:

$$C^* = \underset{C}{\operatorname{argmin}} E(C) = \frac{1}{2} \int_0^1 [\alpha |C'(p)|^2 + \beta |C''(p)|^2 - E_{ext}(C(p))] dp, \quad (2.1)$$

where

$$E_{ext}(x, y) = \|\nabla[G_\sigma(x, y) * I(x, y)]\|^2,$$

and C' and C'' are the first and second order derivatives of the curve, and $p \in [0, 1)$ is an arbitrary parametrization, $\alpha, \beta > 0$ are parameters to balance the effect of each terms. Minimizing the integration of $|C'(p)|^2$ will provide us with a smooth curve of small first order derivative. Minimization of the integration of $|C''(p)|^2$ can provide us with a curve whose first-order derivative is also smooth. These two terms are from prior knowledge of the boundary of an object. The last term is a data term corresponding to the image gradient.

The Geodesic Active Contour (GAC) model [6, 17] interprets the Snakes as a

Riemannian shorted path problem. The GAC is formulated below:

$$C^* = \operatorname{argmin}_C \int_0^L g(C(p)) \|C_p\| dp, \quad (2.2)$$

where

$$g(x, y) = \frac{1}{1 + \|\nabla G_\sigma * I\|} \quad (2.3)$$

is the Riemannian metric, G_σ is a 2D Gaussian filter in a scale determined by σ , and L is the bound for the parametrization of p . The function g is an edge indicator function, a.k.a. an edge map.

2.2.2 Region-based active contour models

Edge-based active contours are known to be sensitive to initialization and noise. Region-based active contours have been proposed as an alternative to the edge-based active contours. The first region-based active contour is the *region competition* [18]. The region competition is actually a realization of the simplest case of the framework of energy minimization originally proposed by Mumford and Shah [20, 37], which is now known as the Mumford-Shah model.

In the original Mumford-Shah model, images are considered to be piecewise smooth. The image $I(x, y)$ on the image domain Ω is modeled via a piecewise smooth function $f(x, y)$ which is allowed to be non-smooth at the boundary of the segmented disjoint smooth regions, and $f(x, y)$ is as smooth as possible within each smooth regions. The segmentation is modeled as an energy minimization problem as follows:

$$\min_{f, \Gamma} E(f, \Gamma) = \int_{\Omega} (f - I)^2 dx dy + \lambda \int_{\Omega - \Gamma} \|\nabla f\|^2 dx dy + v |\Gamma|, \quad (2.4)$$

where Γ is the boundary, λ and v are the penalty parameters. This model is often

referred to as the full Mumford-Shah functional, which is difficult to be minimized.

The region competition [18] is based on a simplified form of the Mumford-Shah model, i.e. the Mumford-Shah cartoon model [20], it is assumed that the image $I(x, y)$ defined over the image domain Ω can be approximated via piecewise constants $c_l, l = 1, 2, \dots, n$ on the disjoint regions (segments) $\Omega_l, l = 1, 2, \dots, n$, where $\Omega = \bigcup_l \Omega_l$. The energy minimization model of this idea can be written as follows:

$$\min_{\{c_l\}, \{\Omega_l\}} E(\{c_l\}, \Gamma) = \sum_{l=1}^n \int_{\Omega_l} (c_l - I)^2 dx dy + v|\Gamma|. \quad (2.5)$$

In fact, it is proven in [20] that the full model as expressed in Eq. (2.4) tends to be the cartoon model when increasing the penalty of smoothness λ .

A number of research works on the two-phase Mumford-Shah model have been reported, such as the famous reformulation by Chan and Vese [24]. For an image defined over the image domain Ω , the two-phase Mumford-Shah model can be formulated as follows:

$$\begin{aligned} \min_{\phi} E(\phi, g_1, g_2) = & \int_{\Omega} \mathcal{E}_1(x, y, g_1) H(\phi) dx dy \\ & + \int_{\Omega} \mathcal{E}_2(x, y, g_2) (1 - H(\phi)) dx dy \\ & + \nu \int_{\Omega} \delta(\phi(x, y)) \|\nabla \phi\| dx dy, \end{aligned} \quad (2.6)$$

where $\delta(\cdot)$ is a Dirac delta function, $H(\cdot)$ is a Heaviside function, ν is a penalty coefficient, g_1 and g_2 are the reconstruction estimates in the two image regions, and ϕ is a signed distance function which can be defined as follows:

$$\phi(x, y, t) = \begin{cases} d([x, y]^T, \partial\Omega), & [x, y]^T \in \Omega \\ 0, & [x, y]^T \in \partial\Omega \\ -d([x, y]^T, \partial\Omega), & [x, y]^T \notin \Omega, \end{cases} \quad (2.7)$$

where Ω is the object region and $\partial\Omega$ is the boundary of Ω . The signed distance function represents the partition of the image, i.e. $\phi > 0$ for one region, $\phi \leq 0$ for the other, and it also implicitly represent the boundary contours of the two regions. In Eq. (2.6), $\mathcal{E}_1$ and $\mathcal{E}_2$ are the reconstruction errors of the two region models. H assigns either of the two models to all the pixels. The last term tries to minimize the geometrical complexity of the labeling, and it has been interpreted as a Minimum Description Length principle by Zhu and Yuille [18].

If we assume the reconstruction functions g_1, g_2 to be constants, i.e. for all (x, y) in Ω , $g_1(x, y) = c_1$ and $g_2(x, y) = c_2$, then the gradient terms in the original model vanish, and the functional becomes the piecewise constant (PC) Mumford-Shah model, which is the prototype of the Region Competition [18] and the Chan-Vese model [24]. For the piecewise smooth (PS) model, we have $\mathcal{E}_1 = (I - g_1)^2 + \lambda \|\nabla g_1\|^2$, $\mathcal{E}_2 = (I - g_2)^2 + \lambda \|\nabla g_2\|^2$.

It may be noticed that the edge-based active contour energies, e.g. the Snakes model in Eq. (2.1) and GAC model in Eq. (2.2), are formulated as contour integrals, whereas the region-based active contour energies, e.g. in the Mumford-Shah model, are formulated as region integrals. It might be interesting to know why the two distinct types of formulations fall into the same category, i.e. they are all active contour models. Perhaps, the first work on connecting the two types of models is the region competition [18]. In that seminal work, the authors established the equivalence between the region integral of the PC Mumford-Shah model and a contour integral by using the Green's theorem. A detailed derivation of the region competition is given in Appendix A.

2.3 A unified view of edge-based and region-based active contour models

In the following, it is shown that the active contour models in the original forms of contour integral and region integral can be unified as region integral in terms of signed distance functions. The connection is based on the co-area formula in 2D [38] given as follows:

$$\iint_{\mathcal{D}} g \|\nabla u\| dx dy = \int_{-\infty}^{\infty} \int_{\{\Gamma(s)|u=t\}} g ds dt. \quad (2.8)$$

Let us consider the GAC energy in Eq. (2.2). We may try to rewrite the GAC energy as an region integral in terms of signed distance functions as follows:

$$\begin{aligned} \int_0^L g \|C_p\| dp &= \int_{\{C(s)|\phi=0\}} g ds \\ &= \int_{-\infty}^{\infty} \int_{\{C(s)|\phi=t\}} g \delta(t) ds dt \\ &= \iint_{\mathcal{D}} g \delta(\phi) \|\nabla \phi\| dx dy, \end{aligned} \quad (2.9)$$

where $\delta(t) = \delta(\phi)$ at all the locations, and $\delta(\cdot)$ is a Dirac delta function.

The *Gâteaux* derivative of the GAC energy in the region integral form is the following:

$$\begin{aligned} &\frac{d}{d\epsilon} \iint_{\mathcal{D}} g \delta(\phi + \epsilon\psi) \|\nabla \phi + \epsilon \nabla \psi\| dx dy \\ &= \iint_{\mathcal{D}} g \frac{d\delta(\phi)}{d\phi} \psi \|\nabla \phi\| dx dy + \iint_{\mathcal{D}} g \delta(\phi) \frac{\nabla \phi \cdot \nabla \psi}{\|\nabla \phi\|} dx dy \\ &= - \iint_{\mathcal{D}} \left\langle \nabla g, \frac{\nabla \phi}{\|\nabla \phi\|} \right\rangle \delta(\phi) \psi dx dy - \iint_{\mathcal{D}} g \nabla \cdot \left(\frac{\nabla \phi}{\|\nabla \phi\|} \right) \delta(\phi) \psi dx dy. \end{aligned} \quad (2.10)$$

The derivation is similar to Eq. (A.31).

The gradient descent equation of the energy is

$$\frac{\partial \phi}{\partial t} = \left\langle \nabla g, \frac{\nabla \phi}{\|\nabla \phi\|} \right\rangle \delta(\phi) g + \nabla \cdot \left(\frac{\nabla \phi}{\|\nabla \phi\|} \right) \delta(\phi). \quad (2.11)$$

Note that $\frac{\nabla \phi}{\|\nabla \phi\|}$ at $\phi = 0$ is the normal of the zero level set, $\nabla \cdot \left(\frac{\nabla \phi}{\|\nabla \phi\|} \right)$ is the curvature. The above is essentially the level set form of the curve evolution of GAC in (A.17).

We may combine the above region integral with that of the Region Competition to arrive at the following general active contour model:

$$\Omega^* = \operatorname{argmin}_{\Omega} \mathcal{G}(\Omega), \quad (2.12)$$

and

$$\begin{aligned} \mathcal{G}(\Omega) = & \gamma_1 \int_{\mathcal{D}} \hat{g}(x, y) \delta(\phi_{\Omega}(x, y)) dx dy \\ & + \gamma_2 \int_{\Omega} g(x, y) dx dy + \gamma_3 \int_{\mathcal{D} \setminus \Omega} \bar{g}(x, y) dx dy, \end{aligned} \quad (2.13)$$

where $\Omega \in \mathcal{D}$ is the region to optimize, $\mathcal{D}$ is the entire integral domain, $\hat{g}(x, y)$, $g(x, y)$ and $\bar{g}(x, y)$ are function mappings defined over the entire domain, $\hat{g}(x, y)$ is a measure of boundary optimality, and $g(x, y)$ and $\bar{g}(x, y)$ measure the region optimality. $\phi_{\Omega}(x, y)$ is a signed distance function defined in Eq. (2.7) in which Ω is the region defined by the closed contours C , such that C is the boundary of Ω . Besides, $\delta(\cdot)$ is a Dirac delta function, and $\gamma_1, \gamma_2, \gamma_3$ are penalty coefficients.

Together with the Chan-Vese model, we may rewrite Eq. (2.12) as follows:

$$\phi_{\Omega}^* = \operatorname{argmin}_{\phi_{\Omega}} \mathcal{G}(\phi_{\Omega}), \quad (2.14)$$

and

$$\begin{aligned}
 \mathcal{G}(\phi_\Omega) &= \gamma_1 \int_{\mathcal{D}} \hat{g}(x, y) \delta(\phi_\Omega) dx dy \\
 &\quad + \gamma_2 \int_{\mathcal{D}} g(x, y) H(\phi_\Omega) dx dy \\
 &\quad + \gamma_3 \int_{\mathcal{D}} \bar{g}(x, y) (1 - H(\phi_\Omega)) dx dy,
 \end{aligned} \tag{2.15}$$

where $H(\cdot)$ is a 1-D Heaviside function.

Chapter 3

Active Contours with a Tangential Component

The energy-minimizing curve evolution of edge-based active contours can terminate early before converging to the object boundaries of interest. This chapter presents the investigation of the cause of this problem. A solution to this problem is proposed based on the observations presented in this chapter. It is observed that the full gradients of the edge indicator along the boundaries are often uniformly small, i.e. both the normal and tangential components of the gradients of the edge indicator along the boundaries are close to zero. However, conventional edge-based active contours, such as the Geodesic Active Contour (GAC), often converge to the contour curves on which the tangential component is still significant. This chapter subsequently presents a new active contour model termed the Geodesic Snakes (GeoSnakes) for locating object boundaries. The Euler-Lagrange (EL) equation of the GeoSnakes requires the full gradients of the edge indicator to vanish at the solution. However, the curve evolution driven by gradient descent in the existing formulations can only solve the normal component of the EL equation, leading to a problem termed the Pseudo Stationary Phenomenon in this chapter. An auxiliary

curve evolution equation, termed the Equilibrium flow (EF) equation, is proposed to specifically solve the tangential component of the EL equation. The full EL equation involving both the normal and tangential components can be solved by alternating the direct gradient descent driven curve evolution and the EF driven curve evolution. Experimental results validate the proposed geometrical interpretation of the early termination problem, and they also show that the proposed method is able to overcome the problem.

3.1 Introduction

Energy minimization provides a principled framework for solving various fundamental computer vision problems. Active contour is proposed as a general framework for object segmentation based on energy minimization. The essential idea of the active contour is to model the object boundaries by the contour curves that minimize the functional energy which measures the error of boundary detection. The active contour has been adopted in many application domains of computer vision, such as surveillance video analysis [8, 36], medical image analysis [6, 32], etc.

Kass et al. [3] proposed the first active contour, commonly known as the Snakes, based on the regularization theory of computational vision [2] and Marr's theory of edge detection [4]. It replaces the traditional heuristic seed growing, region merge/split strategies with the energy-minimizing curve evolution approach. However, the original Snakes requires explicit reparametrization of the contours. Meanwhile, in the area of computational physics, Osher and Sethian [14] proposed the implicit Level Set method for simulating front propagation, which is numerically more advantageous than the conventional explicit parametric schemes used in the original Snakes. Caselles et al. [15] and Malladi and Sethian [16] derived new

curve evolution equations that can be solved by the level set method for boundary extraction without incorporating active contour energies. Subsequently, Caselles et al. [17] and Yezzi et al. [6] separately obtained the geometrical formulation of the active contours, which leads to the new framework of energy minimization by Level Set method.

There are both edge-based and region-based active contours [18, 24, 28, 29]. It is arguable that in general neither the region-based nor the edge-based methods can outperform the other. There are some recent attempts on integrating the region-based and edge-based methods [39, 40], improving the region-based methods [41], and improving the edge-based active contours [42, 43]. The work in this chapter is devoted to the improvement of the general edge-based active contours. Specifically, this chapter addresses a problem that was reported in [5, 42–45], which was not yet explored in depth. The problem was often stated as: *when the curve is initialized relatively arbitrarily, the active contour can stop early and some part of the converged curve can still be far from the boundaries of the objects of interest.* The relatively arbitrary initialization is desired in the practice of semi-automatic object segmentation.

For boundary extraction, the object boundaries perceived by human are located at the peaks of the magnitude of image gradients according to the theory of edge detection [4]. Based on the magnitude of image gradients, an edge indicator function can be formulated as in Eq. (2.3). Variants of the edge indicator functions have now been commonly used in the formulation of active contours. A prototypical edge-based active contour is the Geodesic Active Contour (GAC). The Euler-Lagrange (EL) equation for minimizing the GAC energy requires the gradients of the edge indicator in the normal direction of the curve to approach zero. However, it has been observed that the curve evolution for solving the EL equation often converges at the places where the tangential component is still sig-

Figure 3.1: The full gradients of the edge indicator near the object boundary are close to zero. (a) is an image containing an object of interest. In (b), the center is the edge indicator function for (a). The intensity of the edge indicator is visualized by the darkness in figure (b). The surrounding patches in (b) show the gradient vector field of the edge indicator along the object boundary (dashed curve).

nificant [5, 42]. In fact, the *full* gradients of the edge indicator along the boundaries are often small, i.e. both the normal and tangential components of the gradients of the edge indicator along the boundaries are close to zero, as illustrated in Figure 3.1. Ensuring only the normal component of the gradients to be zero is therefore insufficient for boundary detection. This problem with GAC may be considered as a problem of local optimality of the curve evolution based method, but the globally optimal solution of GAC can be a dot, because GAC tries to find the contour that has the minimal weighted contour length.

This chapter presents a novel active contour model, the Geodesic Snakes (GeoSnakes) model. The derived Euler-Lagrange (EL) equation of the GeoSnakes model requires the full gradients of the edge indicator to be close to zero for the convergence of the curve evolution. In other words, the EL equation of this Snakes model is the EL equation of GAC plus one additional tangential component. This is in contrast to the conventional curve evolution method which only solves for the normal component in the EL equation and the curve evolution would stop even if the

tangential component of the EL equation is not fully solved. This phenomenon is termed as the Pseudo Stationary Phenomenon (PSP) of the curve evolution. Therefore, the curve evolution for solving the GeoSnakes may not achieve the minimal GeoSnakes energy because of the PSP, but it minimizes the GAC energy. The above discussions give rise to a mathematical interpretation of the early termination of the curve evolution by GAC. To address the PSP, an auxiliary curve evolution equation, named the Equilibrium Flow (EF), is proposed accordingly to solve specifically for the tangential component of the EL equation of the GeoSnakes model. The full EL equation of the GeoSnakes for boundary extraction is solved by applying the Equilibrium Flow and the GAC flow in turns.

The contributions of this chapter are listed as follows:

- (i) it has been elucidated that, in addition to the normal component, the tangential component of the gradient of the edge indicator function along the object boundary, is also important for boundary extraction;
- (ii) a new active contour model is proposed. The EL equation of this new active contour requires both the normal and tangential component of the gradient of the edge indicator function along the contour to approach zero; and
- (iii) a fundamental solution to the EL equations containing both the normal and tangential components is proposed.

The rest of this chapter is organized as follows. Section 3.2 introduces the related works addressing the problem of early termination of curve evolution. Section 3.3 presents the problem statement followed by the proposed formulation and solution accompanied with the associated theoretical analysis. Section 3.4 presents the experiments conducted to validate the interpretation of the early termination problem and evaluate the proposed method. Section 3.5 concludes the chapter with suggestions on possible future work.

3.2 Previous works on the early termination problem in curve evolution

The early termination of edge-based active contours was first reported in [44] by Cohen. He proposed the Balloon term to push the curve either to shrink or to expand. Generally, the Balloons can be applied when there is little attraction sensed by the evolving curve. Xu and Prince [5] observed that the Snakes might not converge to concave boundary even when there is large attraction. Their statement is quoted here:

Although the external forces¹ correctly point toward the object boundary, within the boundary concavity the forces point horizontally in opposite directions. Therefore, the active contour is pulled apart toward each of the fingers of the U-shape, but not made to progress downward into the concavity

A U-shape as well as the associated normalized velocity field is reproduced and shown in Figure 3.2 to visualize the statement above. Based on this observation, Xu and Prince [5] proposed the Gradient Vector Flow (GVF) to address the problem of extracting objects with moderately concave boundaries. This method is tested to be effective for moderate boundary concavity.

Paragios et al. [42] attempted to apply the GVF to GAC where the Level Set method is used for extracting multiple objects. However, they discovered a problem which is quoted here:

...the proposed flow does not perform propagation when the NGVF² is close to orthogonal to the inward normal ... propagation will not take

---

¹the gradient vectors

²normalized gradient vector flow

Figure 3.2: The U-shape and the opposite gradient vectors

place, as well as change of the topology, even if they are supported by the level set technique.

We can compare this statement with the observations by Xu and Prince [5] and conclude that they reported the same problem. Paragios and his colleagues finally suggested using region based active contours such as Geodesic Active Regions [46] instead of edge (local gradient) based active contours.

Li et al. [45] proposed to pre-segment the image before starting curve evolution. The segmentation is based on the geometry of the GVF. By segmenting the image, the closed curve can then be cut into smaller ones. However, this work redirects us to another area, image segmentation based on GVF.

More recently, Xie and Mirmehdi [43] formulated a novel edge-based curve evolution equation motivated by the mathematical formulation of magnetostatic/Lorentz force, as an alternative to the conventional edge-based active contours. This curve evolution model was thus named the Magnetostatic Active Contour (MAC). In that paper, the authors conjectured that the early termination of edge-based curve evolution can be due to the undesired critical points, such as saddle points or maxima. The authors also demonstrated that MAC can surpass the undesired critical points and converge to boundaries with severe concavities or multiple objects for many images while having less constraints on the initializations. However, MAC requires

the edge detection to produce little spurious edges as a preprocessing step. Besides, the detected edge helps produce an indicator of region homogeneity, and the resulting curve evolution behaves like region based active contours as observed in the experiments presented subsequently in this chapter. Moreover, MAC is a curve evolution framework, and the curve evolution does not necessarily minimize an energy functional. Hence, MAC is not formulated under the energy minimization framework.

3.3 The Geodesic Snakes model

3.3.1 Problem statement

Figure 3.3 shows a possible situation of early termination of the curve evolution in GAC. Let the curve segments at A and B be the same curve segment from two successive iterations of the curve evolution in a vector field. The black arrows in Figure 3.3(b) denote the actual attraction velocity on the contour curves. The attraction velocity is the projection of the vector field, visualized by the red arrows, in the normal direction of the curve. As the curve evolves, the curve at A moves towards position B in the first iteration. Then, in the second iteration, the curve at B will move towards A . By using a sufficiently small time step, the curve will finally stop at some place where the gradients are nearly orthogonal to the inward/outward normal, but not necessarily zero. The curve evolution terminates too early since the curve converges to the place in-between the boundaries where the normal components vanishes but not the tangential ones. All these observations imply that the *tangential component of the gradients of the edge indicator is also important for object boundary extraction.*

Figure 3.3: The curve stops (or oscillates) when the gradient is nearly orthogonal to the inward normal of the curve. (a) is a 3D visualization of a typical edge indicator function overlaid by the negative gradient field; (b) demonstrates the curve evolution in the negative gradient field within the gray-shaded box region in (a).

3.3.2 The GeoSnakes and the Pseudo Stationary Phenomenon

In this subsection, a new active contour model that requires the full gradients of an edge indicator to approach zero at the optimal solution is proposed. In essence, the EL equation of this active contour model is the EL equation of the GAC plus an additional tangential term. The model is given as follows:

$$C^* = \operatorname{argmin}_C \mathcal{L}(C) = \int_C (\ln \circ g \circ C) dp + \int_C \|C_p\| dp, \quad (3.1)$$

where $\circ$ is the function composition operator, $\ln \circ g \circ C = \ln(g(C))$ is the composition of the three functions, and g is the edge indicator function defined in Eq. (2.3). The corresponding EL equation is the following:

$$\begin{aligned} \frac{\delta \mathcal{L}(C)}{\delta C} &= \frac{1}{g} \nabla g - \kappa \vec{N} = 0 \Leftrightarrow \nabla g - g \kappa \vec{N} = 0 \\ &\Leftrightarrow \langle \nabla g, \vec{T} \rangle \vec{T} + \langle \nabla g, \vec{N} \rangle \vec{N} - g \kappa \vec{N} = 0, \end{aligned} \quad (3.2)$$

where $\vec{T}$ is the tangential component of the contour curve, and $\vec{N}$ is the normal component. The first equivalence holds since g is guaranteed positive. Comparing Eq. (3.2) with the curve evolution equation of GAC in Eq. (A.17), one can

observe that Eq. (3.2) contains an additional tangential component, which is also the tangential component of the gradient of the edge indicator.

There exist geometric and parametric active contours. A discussion regarding the relationship between the geometric active contours and the parametric active contours is in [47]. The major difference between the two kinds of active contours is that the curve evolution equation of parametric active contours may contain a tangential component whereas the other one contains only the normal component. GAC is a typical geometric active contour. A typical parametric active contour containing tangential component is the original Snakes active contour [3, 47]. The proposed model is named the *Geodesic Snakes* (*GeoSnakes*) since it is related to both the GAC and Snakes.

The corresponding curve evolution equation is as follows:

$$\frac{\partial C}{\partial t} = g\kappa\vec{N} - \langle \nabla g, \vec{T} \rangle \vec{T} - \langle \nabla g, \vec{N} \rangle \vec{N}. \quad (3.3)$$

Now let us turn to the solution of the EL equation of the GeoSnakes model. It is well-known that the tangential component in the curve evolution equations, such as Eq. (3.2), will only automatically reparameterize the contour curve, and the tangential component is conventionally omitted in the implementation of the curve evolution. This behavior can be explained with the help of the Lemma of curve evolution stated in [48].

Lemma 3.1 *Given the closed curve $C(\tilde{p}, t)$ parameterized by arbitrary $\tilde{p} \in \tilde{B}$ at an artificial time t with the normal $\vec{N}$, the tangent $\vec{T}$ of the curve, and given the geometric flow of a curve evolution by*

$$\frac{\partial C(\tilde{p}, t)}{\partial t} = \alpha(\tilde{p}, t)\vec{T}(C(\tilde{p}, t)) + \beta(\tilde{p}, t)\vec{N}(C(\tilde{p}, t)), \quad (3.4)$$

and if β does not depend on the parametrization, meaning that β is a geometric

intrinsic characteristic of the curve, then the image of $C(\tilde{p}, t)$ that satisfies Eq. (3.4) is identical to the image of the family of curves $C(p, t)$, parameterized by $p \in B$, that satisfies

$$\begin{cases} \frac{\partial C(p, t)}{\partial t} = \beta(p, t) \vec{N}(p, t) \\ \frac{\partial p}{\partial t} = -\frac{\alpha}{\|C_p\|}. \end{cases} \quad (3.5a)$$

$$(3.5b)$$

The above lemma tells us that the geometry of the curve evolution only depends on the velocity in the normal direction of the curve, although the reparametrization by $\frac{\partial p}{\partial t}$ may last forever due to the possibly non-vanishing α . A brief proof of this lemma can be found in the book written by Sapiro [11].

The omission of the tangential component in the original EL equation can also be seen in the level set method for implementing general curve evolution. Recall the level set equation for curve evolution as follows [14]:

$$\begin{aligned} \left. \frac{\partial \Phi}{\partial t} \right|_{\Phi=0} &= \langle \nabla \Phi, \partial_t C \rangle|_{\Phi=0} \\ &= \langle \nabla \Phi, \beta \vec{N} + \alpha \vec{T} \rangle|_{\Phi=0}, \text{ by Equation(3.4)} \\ &= \beta \|\nabla \Phi\| \Big|_{\Phi=0} \\ &\text{by setting } \vec{N} = \frac{\nabla \Phi}{\|\nabla \Phi\|} \Rightarrow \langle \nabla \Phi, \vec{T} \rangle = 0. \end{aligned} \quad (3.6)$$

From Eq. (3.6), the consistency between the level set method and the geometric behavior of curve evolution described in Lemma 3.1 is obvious.

All the above arguments conclude that the curve evolution is independent of the tangential velocity on the closed curve. It also implies that the converged solution of Eq. (3.5) may not necessarily be the stationary solution of Eq. (3.4) which we are required to solve. Specifically, there may exists a non-vanishing reparametrization velocity, i.e. the tangential velocity, on the converged curve. This can be ascertained by rephrasing Lemma 3.1 as follows.

Corollary 3.1 (Pseudo Stationary Phenomenon) *The converged curve of Eq. (3.5), denoted by $C(\tilde{p}, \infty)$ which satisfies*

$$\frac{\partial C(\tilde{p}, \infty)}{\partial t} = \beta(C(\tilde{p}, \infty))\vec{N}(C(\tilde{p}, \infty)) = \vec{0} \quad (3.7)$$

can lead to

$$\begin{aligned} \frac{\partial C(p, \infty)}{\partial t} &= \alpha(C(p, \infty))\vec{T}(C(p, \infty)) \\ &\quad + \beta(C(p, \infty))\vec{N}(C(p, \infty)) \\ &\neq \vec{0}, \end{aligned} \quad (3.8)$$

where $p = p(\tilde{p}, t)$, $\alpha = -\|C_p\|\frac{\partial p}{\partial t}$, $\frac{\partial p}{\partial t}$ is the parametrization, and $\alpha(C(p, \infty))$ can be large. The $C(p, \infty)$ refers to the curve at convergence.

The above property is termed the Pseudo Stationary Phenomenon (PSP), since the converged curve may not be a solution of the original EL equation. Therefore, it is technically difficult to solve Eq. (3.2) and to implement the curve evolution defined by Eq. (3.3).

3.3.3 The Equilibrium Flow and the Alternating Curve Evolutions

In the following, a solution to the curve evolution equation with a tangential component, i.e., a solution to the PSP problem, is given. The curve evolution equation concerned can be written in the most general form as follows:

$$\frac{\partial C}{\partial t} = \beta(p, t)\vec{N}(p, t) + \langle \vec{F}(p, t), \vec{T}(p, t) \rangle \vec{T}(p, t), \quad (3.9)$$

where $\vec{F}$ is a vector field. The objective is to ensure that the inner product $\langle \vec{F}(p, t), \vec{T}(p, t) \rangle$ vanishes when the curve evolution stops. Since the curve evo-

lution is independent of the tangential velocity, the tangential component plays no role in the whole process of curve evolution. Therefore, an auxiliary flow termed the *Equilibrium Flow* (EF) is proposed to ensure the tangential velocity to be zero when the curve evolution stops. The EF equation is defined as follows:

$$\frac{\partial C(p, t)}{\partial t} = \langle \vec{F}(p, t), \vec{T}(p, t) \rangle \vec{N}(p, t). \quad (3.10)$$

This equations is named the Equilibrium Flow due to the following propositions.

Proposition 3.1 *The solution C^* of the Equation(3.10) is either on stationary points in $\vec{F}$ or on level sets of the potential g corresponding to the gradient field $\vec{F}$.*

The proof is given in Appendix B.

Another more interesting property of EF is in the following proposition.

Proposition 3.2 *Given the curve evolution of (3.10) and let the reparametrization be $\frac{dp}{dt} = -\frac{\langle \vec{F}(p, t), \vec{N}(p, t) \rangle}{\|C_p\|}$ where $\vec{F} = -\nabla g$, leads to $g(C(p, t)) = g(C(p, t + \tau))$ for any τ before termination of the curve evolution.*

The proof is listed in Appendix B. This proposition implies that the function g for a fixed p would not change during the curve evolution driven by EF. With a proper contour parametrization, the curve evolution driven by EF can solve the tangential component of EL while the energy on the contour will not be changed.

By using this EF flow, a fundamental solution to the EL equation involving both the normal and tangential components is obtained. The formal statement is as follows.

Proposition 3.3 *Given an arbitrary initial curve C_0 , the convergent curve $C(p, \infty)$*

driven by the following system of flows is the stationary solution of Eq. (3.9):

$$\frac{\partial C_1^k}{\partial t} = \beta(p, t) \vec{N}, \quad C_1^k(\tau, 0) = C_2^{k-1}(p, \infty) \quad (3.11a)$$

$$\frac{\partial C_2^k}{\partial t} = \langle \vec{F}, \vec{T} \rangle \vec{N}, \quad C_2^k(p, 0) = C_1^k(p, \infty) \quad (3.11b)$$

$$\begin{cases} k = 1, 2, \dots \\ C_1^0(p, 0) = C_2^0(p, \infty) = C_0 \\ C(p, \infty) = \lim_{k \rightarrow \infty} C_1^k(p, \infty) = \lim_{k \rightarrow \infty} C_2^k(p, \infty) \end{cases} \quad (3.11c)$$

This claim is valid since Eq. (3.9) can be satisfied only when both the β and α are zero, and the curve evolutions defined in Eqs. (3.11a) and (3.11b) ensures the β and α to be zero.

The derived new curve evolution framework can have the following interpretation. Starting from an initial contour curve, the curve evolution of Eq. (3.11a) may converge to a pseudo stationary position because of the PSP. Then the curve evolution of Eq. (3.11b) can help the curve evolution to escape from the pseudo stationary position. At the convergence of Eq. (3.11b), the value of the edge indicator at every contour point of $C(p)$ can be guaranteed unchanged (according to Proposition 3.2), and the curve evolution of (3.11a) has been reactivated because $\langle \nabla g, \vec{N} \rangle$ tends to be large according to Proposition 3.1. The process continues until the system of equations in Eq. (3.11) converges.

The curve evolution of (3.11b) cannot be directly implemented in a level set framework in which the tangent $\vec{T}$ cannot be directly used. Fortunately, in 2D case, there is a useful relation that $\vec{N} = \mathcal{R} \vec{T}$, where $\mathcal{R}$ is a matrix for ± 90 degree rotation. Therefore, the EF can be rewritten as follows:

$$\frac{\partial C(p, t)}{\partial t} = \langle \vec{F}, \vec{T} \rangle \vec{N} = \langle \vec{F}, \mathcal{R} \vec{T} \rangle \vec{N} = \langle \mathcal{R}^T \vec{F}, \vec{N} \rangle \vec{N}. \quad (3.12)$$

Figure 3.4: The practical effect of rotating the gradient field. The top left is a typical gradient field, visualized by ball blue arrows, in which the green line segment cannot be moved by the local gradients due to the PSP. The top right is the vector field generated by the rotation, visualized by the arrows in pinkish red, in which the line segment can now be moved by this vector field. The bottom visualizes both the gradient field and the rotated vector field. It shows that the stationary positions, i.e., the light blue curve, remain stationary after the rotation

The EF Eq. (3.11b) in the system of partial differential equations in Proposition 3.3 can be understood as a repositioning process of the curve. The repositioning process can help the curve escape from the pseudo stationary positions. The visualization of the idea of EF is presented in Figure 3.4. The rotation of the gradient field can reactivate the curve evolution when it is trapped by PSP, while the true stationary positions remain stationary.

The last question we would ask is how the method would behave if the tangential component of the gradient of the edge indicator along the boundary is indeed nonzero, i.e. the EF in Eq. (3.11b) cannot force the contour to converge to the true boundary. Theoretically, the answer to this question relies on the assumption that there exists a stationary solution of Eq. (3.11b) near the true boundary, such that the gradient of the edge indicator in the tangential direction of the contour curve vanishes. Once the curve evolution in Eq. (3.11b) converges to such a contour, the next round of curve evolution by Eq. (3.11a) will converge to the object boundary.

3.3.4 The geometrical analysis of the convergence of curve evolutions

The objective of convergence analysis in this subsection is to check whether the curve evolution obeys the following limiting behavior:

$$\lim_{t \rightarrow t^*} \frac{\partial C(p, t)}{\partial t} = \frac{d\vec{v}}{dt} \Big|_{\vec{v} \in C} = \langle \vec{F}(C, t), \vec{N} \rangle \vec{N} = \vec{0}. \quad (3.13)$$

In the terminology of dynamic system, it is asked whether the convergence of the curve evolution is stable. The answer can be provided by considering a small perturbation to the converged curve, a.k.a. the equilibrium. If the perturbation finally vanishes at the converged curve, the curve evolution is stable, otherwise it is unstable. The stability of the curve evolutions are examined in this manner.

The curve evolution near the converged curve C^* can be represented as follows:

$$C^* = C(p, t^*) = C(p, t)|_{t=t^*-\tau} + \eta(p, \tau), \quad (3.14)$$

where the $\eta(p, \tau)$ is a general formulation of perturbation, $\tau = t^* - t > 0$ and is small enough and monotonically *decreasing* during the evolution. The fundamental conditions of convergence is stated below.

Theorem 3.1 *If $J_{nn}(C^*) < 0$, the convergence of the curve evolution is **stable**. If $J_{nn}(C^*) = 0$, the convergence of the curve evolution is **marginal stable**. If $J_{nn}(C^*) > 0$, the convergence of the curve evolution is **unstable**, in which $J_{nn}(C^*)$ is the projection of the Jacobean of the velocity field $\vec{F}$, denoted by $J[\vec{F}]$, in the normal direction $\vec{N}$ of the converged curve C^* , i.e., $J_{nn}(C^*) = [\vec{N}^T J[\vec{F}] \vec{N}](C^*)$.*

The proof of Theorem 3.1 is established based on several fundamental local geometrical properties of curve evolution. The proof of this theorem and other related theoretical findings are included in Appendix B.

Based on the Theorem 3.1 for examining the stability of convergence of curve evolutions, it is now possible to analyze the proposed system of PDEs in Eqs. (3.11a) and (3.11b).

(i) *For the gradient descent flow, where $\vec{F} = \nabla g$,*

$$\vec{N}^T J[\vec{F}] \vec{N} = \langle \nabla^2 g \vec{N}, \vec{N} \rangle = g_{nn} \quad (3.15)$$

Fortunately, it is possible to perform the analysis of g_{nn} in Lemma B.2 (see Appendix B). The analysis shows that $g_{nn} \leq 0$ on C^* , which proves the following:

Theorem 3.2 *The curve evolution under $\vec{F} = \nabla g$ is stable or at least marginal stable.*

(ii) *For the Equilibrium Flow, where $\vec{F} = \mathcal{R} \nabla g$*

$$\vec{N}^T J[\vec{F}] \vec{N} = \langle \nabla^2 g \vec{N}, \vec{T} \rangle = g_{nt} \quad (3.16)$$

Assume that the solution to EF is C^* , the following equality at C^* holds:

$$\langle \nabla g, \vec{N} \rangle = \langle \nabla g, \pm \frac{\nabla g}{\|\nabla g\|} \rangle = \pm \|\nabla g\|.$$

This is because the gradient vanishes in the tangential direction. Hence the gradient is only in the normal direction. Differentiating the above equality leads to the following:

$$\frac{d}{ds} \pm \|\nabla g(C(s))\| = \pm \frac{\nabla g^T}{\|\nabla g\|} \nabla^2 g C_s = g_{nt},$$

which proves the following.

Theorem 3.3 *If the gradient field is normalized, e.g., $\|\nabla g\| = 1$, the EF will be marginal stable.*

This indicates that the convergence of EF is not satisfactory.

It is easy to show that the convergence of EF can be further improved by the following treatment:

$$\begin{aligned}\frac{\partial C(p, t)}{\partial t} &= \langle A^T \nabla g + \epsilon \nabla g, \vec{N} \rangle \vec{N} \\ &= \langle (A^T + \epsilon I) \nabla g, \vec{N} \rangle \vec{N},\end{aligned}\tag{3.17}$$

where $\epsilon > 0$ is chosen to be a constant. Eq. (3.17) leads to the following theorem.

Theorem 3.4 *By normalizing $\|\nabla g\|$, such that $\|\nabla g\| = 1$, the curve evolution defined by Equation(3.17) is at least marginal stable.*

The above slightly perturbed Equilibrium Flow (3.17) behaves analogous to the original Equilibrium Flow. This is shown by the following fundamental theorem.

Theorem 3.5 *For two evolving curves $C(p, t)$ and $S(p, t)$ and given a fixed p , such that $\frac{\partial S(p, t)}{\partial t} = \frac{\partial C(p, t)}{\partial t} + \vec{G}(p, t)$, and $\frac{\partial C(p, t)}{\partial t}$ is Lipschitz continuous with the Lipschitz constant L , given that $\|\vec{G}(p, t)\| \leq \mu, \mu > 0$, then the following holds:*

$$\begin{aligned}\|C(p, t) - S(p, t)\| &\leq \|C(p, 0) - S(p, 0)\| e^{L(t-t_0)} \\ &\quad + \frac{\mu}{L} \{e^{L(t-t_0)} - 1\}.\end{aligned}$$

Theorem 3.5 provides an upper bound of the difference between two curve evolutions. If the difference is small, the two curve evolutions will behave analogously. The proof is not included here, since the proof of a more detailed and general bound for ordinary differential equations, that are related to Gronwall-Bellman inequality, is presented in Chapter 3 of [49].

Algorithm 1 Alternating Curve Evolutions

Input: Input Image I **Output:** C_k

```

1:  $I_{sm} \Leftarrow I$ , by anisotropic smoothing
2:  $g \Leftarrow \nabla I_{sm}$ , by Eq. (2.3)
3:  $\vec{F} \Leftarrow -g$ , by Gradient Vector Flow (GVF)
4:  $c = 0$ 
5: Initialize  $C_0$ 
6: Set MaximumCycle
7: Set MaximumIteration
8: while  $c \leq \text{MaximumCycle}$  do
9:    $k = 0$ 
10:  while  $k \leq \text{MaximumIteration}$  do
11:    if  $\text{mod}(c, 2) = 0$  then
12:       $C_{k+1} \Leftarrow C_k$  by the direct gradient decent flow of Eq. (3.11a)
13:    else
14:       $C_{k+1} \Leftarrow C_k$  by the revised EF, i.e., Eq. (3.17).
15:    end if
16:     $k = k + 1$ ;
17:  end while
18:   $c = c + 1$ ;
19: end while

```

3.3.5 The curve evolution algorithm for GeoSnakes

To summarize, the proposed method is a curve evolution defined by the system in Eqs. (3.11a) and (3.11b). In practice, Eq. (3.11b) may be replaced by Eq. (3.17). In the implementation of the proposed method, we often require the image to be smoothed before the curve evolution. An edge map is then generated by using the smoothed image according to Eq. (2.3). Simultaneously, an initial contour is to be provided by the user. Afterwards, the curve evolution is run till convergence. The level set method [13] is adopted for implementing the curve evolution. The pseudo code of the method is shown in Algorithm 1.

3.4 Experimental results

The purpose of this section is to validate the claim that the PSP problem is a cause of the early termination problem, and to demonstrate the capability of the proposed method for coping with the PSP problem. The proposed method is compared with other related methods to further assess the performance of the proposed method.

3.4.1 Experimental settings

In this subsection, the parameter settings in the methods used for preprocessing and comparison are presented. The implementation details can be found in the original papers cited below.

Edge map[16, 17] The σ of the Gaussian kernel that is used to generate edge map g in Eq. (2.3) is empirically chosen to be 3 for all images.

GVF[5] The diffusion of the gradients of the edge map g is done by GVF as a preprocessing for all the gradient-based methods compared in the experiment. The GVF is implemented by using the source code provided in the GVF web site³, and $\mu = 0.1$.

Balloon[17, 44] The Balloon is chosen to be positive to shrink the curve.

Adaptive Balloon[42, 50] The version in [50] is adopted, which is simpler for analysis comparing to [42], while they are almost equivalent. In this implementation $\lambda = 1$, $\beta = 0.1$.

Chan-Vese model[24] The parameters λ_1 and λ_2 are set equal to 1, $\mu = 0.1$, and $v = 0$ which is a Balloon term. The approximated delta function δ_ϵ and Heaviside function H_ϵ follow the forms in [24].

---

³<http://iac1.ece.jhu.edu/projects/gvf/>

MAC [43] Canny edge detection is used as a preprocessing, and $\alpha = 0.5$ for all the images.

Approximated Dirac delta The approximated delta function δ_ϵ for level set implementation, excluding Chan-Vese model, is chosen to be:

$$\delta_1(x) = \begin{cases} 0, & |x| > 1 \\ \frac{1}{2}(1 + \cos(\pi x)), & |x| \leq 1 \end{cases}.$$

Re-initialization To guarantee the stability of the curve evolution, the first-order accurate Essentially Non-Oscillatory is adopted instead of the method proposed by Li et al. [51]. We used Baris Sumengen's MATLAB toolbox of the level set method ⁴ for implementation.

Active contour codes The coding of the active contours is largely inspired by Chunming Li's implementation⁵.

The Balloon is chosen to be 1. The algorithms are implemented in MATLAB on a PC.

3.4.2 Segmentation results

Figure 3.5: Segmentation of the U-shape. (a) is the initialization. (b) is the PSP. (f) is the converged result.

⁴http://barissumengen.com/level_set_methods/

⁵<http://www.engr.uconn.edu/~cmli/>

Figure 3.6: Inputs and initializations

Before presenting the main results of the segmentation, the proposed method is demonstrated for segmentation of the U-shape in Figure 3.2. The result is shown in Figure 3.5. The other images used for evaluating the proposed method are shown in Figure 3.6 overlaid with the red initial curves. The same initializations are used in all the comparisons.

The chosen images are of different characteristics. The image of two rectangles in Figure 3.6(a) is relatively easy for segmentation since the gray level of both foreground and background are homogeneous and distinct, the edges are strong too. Besides, there are only two objects, hence the topology is simple. In Figure 3.6(b), the image is also simple except that there are three objects, which is a little more difficult than the former one. In Figure 3.6(c), there are nine curve segments forming three circles. The edge-based active contours may see three circles of disconnected boundaries in the image, and the region-based method may see the objects in the image as composed of nine narrow and curved regions. Figures 3.6(d) and 3.6(e) are magnetic resonance imaging (MRI) medical images of human brain with multiple tumor regions. The original brain scans are shown in Figure 3.7. From Figures 3.6(d) and 3.6(e), we see that there are many curved-edges in the images. Besides, there are some regions with intensities similar to the targeted tumor regions. These make the region-based active contour to be unsuitable for dealing with such images. However, for edge-based active contours, there also exists the Pseudo Stationary Phenomenon.

Figure 3.7: The original real images and the tumor regions (in boxes)

Figure 3.8: Gradient fields, PSP and Saddle points

In Figure 3.8, the *normalized* gradient field of the edge indicator functions of the respective images in Figures 3.6 (a), (b) and (c) are visualized to illustrate the PSP. The saddle points are highlighted by the blue dots on the red lines of which the gradients are in the tangential direction. The effect of the PSP problem on the curve evolution near those lines is studied next.

The curve evolution processes of the system of Eqs. (3.11a) and (3.11b), called the GAC+EF, are shown in Figure 3.9. The objects are correctly segmented. Figure 3.10 shows the segmentation results produced by other methods, including the GAC, GAC with Balloon, GAC with adaptive Balloon, MAC and the Chan-Vese model, with the same curve initialization as shown in Figure 3.6. It can be observed that none of these compared methods can extract all the objects accurately like GAC+EF. GAC suffers from the early stop of the curve evolution. GAC with Balloon can surpass edges and converge to the interior boundaries or vanish; GAC with adaptive Balloon can be applied for extracting two objects

Figure 3.9: Object Extraction by proposed method. The first row is the curve evolution by GAC; The second is the curve evolution by EF; The bottom row is the final convergence.

of simple topology as shown in Figure 3.10, yet it still suffers from the PSP in other images. The gradient field used for the former three methods are extended by Gradient Vector Flow. Interestingly, both the GAC and the GAC with the adaptive Balloon might stop quite close to the pseudo stationary positions as shown in Figure 3.3, while GAC with adaptive Balloon can escape from some pseudo stationary positions of relatively simple topology. It can also be observed that both MAC and the Chan-Vese model can extract the narrow curved regions in the image of three disconnected circles as shown in Figure 3.10. It implies that MAC shares certain characteristics of the region-based methods. The edge detection process in MAC does not produce a local gradient field but the indication of intensity homogeneity. The Chan-Vese active contour is not sensitive to initialization but it often captures other undesired regions. The MAC generally fails to converge to the object boundaries if there are too many edges and corners distributed everywhere in the images.

Figure 3.10: Results by other methods. First column from the left is the boundary extraction by GAC; Second column is the one by GAC with Balloon; Third column is the one by GAC with adaptive Balloon; Fourth column is the one by MAC; Last column is the one by Chan-Vese active contour.

Figure 3.11 visualizes the 3D tumor model that is obtained by the segmentation of the entire MRI sequence of Figure 3.6(e) using GAC+EF. The cutting plate corresponds to the segmentation result shown in the bottom right of Figure 3.9.

Figure 3.11: The reconstructed 3D tumor from the MRI sequence of Figure 3.6(e).

3.4.3 The tangential and normal velocities

In this subsection, a quantitative analysis of the proposed method is provided. The major contribution in this chapter is the proposed solution, namely the EF, to the Pseudo Stationary Phenomenon that occurs if the tangential velocity on the converged curve does not vanish. Thus, it can be demonstrated that the proposed method can reduce the tangential velocity effectively. The tangential velocity in the curve evolutions of the proposed method is shown in the top row of Figure 3.12. The tangential velocity is quite large when GAC converges (0-k1), but due to the EF (k2-k4) the tangential velocity approaches zero, where k3 is the iteration of convergence. The velocities do not reach zero exactly because of the errors in discretization. The normal velocity in the second row of Figure 3.12 shows that the normal velocity drops significantly during the direct curve evolutions of the gradient descent flow. The iterations corresponding to k1, k2, k3, and k4 in Figure 3.12 are given in Table 3.1. The MaximumIteration is 2000 for all the experiments. The computational times for the curve evolutions are also presented. The computational cost can be reduced by using fast implementations, e.g. [52–54].

Figure 3.12: The normal and tangential velocities during curve evolutions. The top row shows the total magnitude of the tangential velocity along the curves; The second row shows the total magnitude of the normal velocity along the curves. k_1, k_2, k_3, k_4 are some key iterations in the curve evolutions. 0- k_2 is the curve evolution by GAC; k_2 - k_4 is by EF; from k_4 onwards is by GAC

3.5 Concluding remarks

The PSP problem may happen not only in object segmentation. It may also exist in other computer vision or computer graphics tasks involving contour optimization. Hence, the proposed method may be extended to address the PSP in different

Table 3.1: Timings for curve evolutions on images of different sizes

Input image	Image size (px)	k1	k2	k3	k4	Total time (s)	GAC time (s)	EF time (s)
	80×80	150	300	900	1800	62.44	4.68	18.73
	80×80	230	300	800	1800	62.33	7.17	15.58
	80×80	200	300	920	1800	63.34	6.33	19.64
	128×128	110	300	1600	1800	119.44	6.57	77.64
	128×128	120	300	850	1800	120.23	7.21	33.06

contexts. The proposed method is derived based on geometrical observations and analysis of the PSP problem. Experimentally, it shows that the EF repositions the curve to escape from the PSP. This suggests that one may also approach the PSP problem from the perspective of repositioning a pseudo-stationary curve.

The problem of early termination of curve evolution for a general class of edge-based active contours was investigated. It is shown that the omittance of the tangential component in the gradient of the edge indicator function may lead to this early termination. Then, a new active contour model called the GeoSnakes was proposed. The EL equation of GeoSnakes requires the full gradient of the edge indicator on the solution to be zero. It can also be noticed that the direct curve evolution by gradient descent for solving the EL equation will ignore the tangential component of the EL equation, leading to the PSP. Hence, an auxiliary curve evolution called the EF is proposed to account for the tangential component. Thus,

---

the full EL equation is solved by alternately applying gradient descent and the EF. This work provides the geometric insight into the problem of early termination of the curve evolution, giving rise to a new edge-based active contour framework.

Chapter 4

Region-based Active Contour for Image Segmentation and Region Description for Non-smooth Image Structures

This chapter presents a unified energy minimization framework for the segmentation of non-smooth structures. The energy of piecewise linear patch reconstruction is proposed to be an objective measure of the quality of the segmentation for non-smooth structures. The segmentation task is accomplished by minimizing the single energy of piecewise linear patch reconstruction without a separate process of feature extraction. It is proven that the error of segmentation is bounded by the error of piecewise linear patch reconstruction, meaning that minimizing the proposed energy leads to reduction of the error of segmentation. As a by-product, the proposed method produces a dictionary of optimized orthonormal descriptors for each segmented region. The experiments validate the theoretical claims and they show the clear superior performance of the proposed method over other related

methods for segmentation of various image textures containing rich non-smooth image structures. It is also shown that the proposed model could be coupled with the piecewise smooth model to handle both smooth and non-smooth image structures. It is demonstrated that the proposed model is capable of coping with multiple regions with different structures through the one-against-all strategy.

4.1 Introduction

The quality of the solutions to image segmentation problems should be measured objectively using mathematical models. An objective measure of piecewise smooth image segmentation is the Mumford-Shah functional energy [20]. By minimizing the energy, we expect to achieve the segmentation of good quality for piecewise smooth images. The problem of minimization of the Mumford-Shah functional energy is the mathematical abstraction, i.e. the modeling, of piecewise smooth image segmentation. The model is now known as the Mumford-Shah model.

In the Mumford-Shah model, each image region is modeled as a smooth or constant function with parameters. By minimizing the energy of the Mumford-Shah model, we can restore the smooth image in each region, and we can also obtain the partition boundaries of the segmentation located at the discontinuities in the restored image. However, the image values to be restored may not be piecewise smooth or flat as a whole. Imposing the smoothness leads to the destructive averaging of the image content. This poses problems to the segmentation with the conventional Mumford-Shah model. For example, it is possible that non-smooth visual patterns different in structure may have similar average image values (see Figure 4.1a). Consequently, the Mumford-Shah model cannot separate such patterns in the image space.

To cope with non-smooth image structure, the segmentation has been consid-

Figure 4.1: Illustration of the problem. The left two images are composed of different image structures that the conventional Mumford-Shah model would fail to segment; The diagram on the right shows that the Gabor feature image of a texture is non-smooth.

ered under a framework of two independent processes, i.e. feature extraction and the segmentation of the feature image. The Mumford-Shah model has been applied to the feature image, such as in the methods of [39, 55], and the more sophisticated framework of [56]. There exist principles for selecting the features from the outputs of filtering without supervision, such as the Dominant Component Analysis [56]. The major problem concerns us is that these frameworks do not provide a unified objective measure of the quality of the segmentation of non-smooth image structures. Thus, the quality of the results can only be assessed by humans eventually. Besides, the combination of the Mumford-Shah model or other active contour models with feature extraction, including [57, 58], is heuristic. Hence, there could be cases in which we should not apply the segmentation to the feature images. For example, in Figure 4.1b the feature values obtained by image filtering with the selected Gabor filter or the local averages of the feature values, can have significant spatial variations. Note that the selected Gabor filter for Figure 4.1b visually matches the structure of the texture. These concerns give us the motivation in this thesis to explore a unified and valid mathematical model for segmentation of non-smooth image structures.

In this chapter, a novel unified energy minimization model for the segmentation of general non-smooth image structures is proposed, which is motivated

by the original Mumford-Shah model. The segmentation of general non-smooth image structures is formulated as a single energy minimization problem of piecewise linear patch reconstruction. The formulation is not based on heuristics. It is proven that the error of segmentation is bounded by the energy of piecewise linear patch reconstruction. Thus, minimizing the energy of piecewise linear patch reconstruction can reduce the error of segmentation. Furthermore, it is also proven that the eigen-patches constitute the global optimal solution to the minimization of the error of linear patch reconstruction. A more efficient alternative to the eigen-decomposition of matrix for computing eigenvectors is also proposed. It is proven that the linear patch reconstruction based on gradient descent converges to the eigen-patches subject to some mild assumption on the initializations. The assumption can often be met, and the convergence is linear. Therefore, the gradient descent algorithm is proposed for solving the linear patch reconstruction problem even though the problem is not convex. The resultant segmentation algorithm is based on alternating piecewise linear patch reconstruction and curve evolution. The piecewise linear patch reconstruction can also be naturally coupled with the conventional Mumford-Shah model for coping with both smooth and non-smooth image structures.

The rest of this chapter is organized as follows. The related works are reviewed in Section 4.2. The mathematical analyses of the piecewise constant or smooth model and the linear patch reconstruction are presented in Sections 4.3-4.4. The error bound of the segmentation in terms of reconstruction error is presented in Section 4.4.2. The theoretical claims of the global optimality of the gradient descent for the nonconvex linear patch reconstruction problem are presented in Section 4.4.3. Experimental results are presented in Section 4.5. Concluding remarks are given in Section 4.7.

4.2 Region-based active contours for unsupervised segmentation of non-smooth image structures

As introduced in the previous section, the original Mumford-Shah model was formulated for segmentation of piecewise smooth images. Subsequently, the model was extended to image texture segmentation, a case of segmenting non-smooth image structures. In [39, 55, 59], Gabor filtering was incorporated in the Mumford-Shah model. Lee et al. [59] embedded the manually selected 24 Gabor filters into the piecewise smooth (PS) Mumford-Shah model. Sandberg et al. [55] adopted the Chan-Vese model, which is the piecewise constant (PC) Mumford-Shah model, as well as the maximum difference of feature means as the criterion for Gabor filter selection. Sagiv *et al.* [39] adopted the framework in [55] with manually selected filters. These frameworks require the textures to be piecewise smooth or flat in the feature space, but the filter selection for unsupervised segmentation cannot ensure this. More recently, Kokkinos *et al.* [56] proposed to apply the Chan-Vese model to a type of modulation features. This framework assumes that the textures are globally oscillating, as on zebras and tigers, and it still requires filter selection by dominant component analysis for parametric texture modeling. However, it is not known *a priori* what kind of feature or the principle of feature selection can help the segmentation by Mumford-Shah model. The existing ways of combining the feature extraction and the Mumford-Shah model do not follow a unified mathematical model for the segmentation.

There exist other active contour frameworks for texture segmentation with pre-selected fixed texture feature [57, 58]. The region competition [18], which is a variant of the piecewise constant Mumford-Shah model, was adopted. In these works, the structure tensor was chosen as the texture feature under an assump-

tion that this texture feature is discriminative across different textures. However, the choice of the structure tensor feature has also not been validated for general textures.

In summary, the existing active contour frameworks for segmentation of non-smooth image structures were not developed under a unified mathematical model of the segmentation but based on heuristics.

4.3 The principle of segmentation behind the Mumford-Shah model

This section covers the study of the segmentation by the Mumford-Shah model with an intention to clarify its rationale. The two-phase model is discussed. The generalization of the two-phase framework to multi-phase may follow [60], which is out of the scope here.

Let us consider the two-phase Mumford-Shah model in a simplified form as follows:

$$H^* = \operatorname{argmin}_H \int_{\Omega} \mathcal{E}_1 H dx dy + \int_{\Omega} \mathcal{E}_2 (1 - H) dx dy, \quad (4.1)$$

where $\mathcal{E}_1$ and $\mathcal{E}_2$ are the reconstruction errors corresponding to the subregions Ω_1 and Ω_2 , such that $\Omega = \Omega_1 \cup \Omega_2$. Note that the prior term of the arc length is omitted, because the term is for imposing smoothness on the boundary. In the Piecewise Constant (PC) model, $\mathcal{E}_1(x, y) = \|I(x, y) - c_1\|^2$, $\mathcal{E}_2(x, y) = \|I(x, y) - c_2\|^2$, where $I(x, y)$ is the image value, and c_1, c_2 are the regional means of the image value. In the Piecewise Smooth (PS) model, $\mathcal{E}_1(x, y) = (I - g_1)^2 + \lambda \|\nabla g_1\|^2$, $\mathcal{E}_2(x, y) = (I - g_2)^2 + \lambda \|\nabla g_2\|^2$, where g_1, g_2 are the smooth restoration of the two image regions (due to the smoothness constraint). The smoothness regularization terms,

$\lambda\|\nabla g_1\|^2$ and $\lambda\|\nabla g_2\|^2$ are often omitted if g_1 and g_2 are solved by the normalized Gaussian convolution [61], which was interpreted as nonparametric regression [28]. Regarding the minimization of this simplified functional in Eq. (4.1), we have the following interesting fact upon fixing the error functions $\mathcal{E}_1(x, y)$ and $\mathcal{E}_2(x, y)$.

Proposition 4.4

$$\begin{aligned} \min_H \int_{\Omega} \mathcal{E}_1 H + \mathcal{E}_2 (1 - H) dx dy \\ = \int_{\Omega} \min_H \{(\mathcal{E}_1 - \mathcal{E}_2) H\} dx dy + C, \end{aligned} \tag{4.2}$$

where $H(x, y) = \{0, 1\}$ and C is a constant independent of H .

This fact tells that optimizing the global assignment of H is equivalent to optimizing the assignment locally. The assignment rule for determining the optimal H at each pixel location is therefore the following:

$$H(x, y) = \begin{cases} 1, & \mathcal{E}_1(x, y) \leq \mathcal{E}_2(x, y) \\ 0, & \mathcal{E}_1(x, y) \geq \mathcal{E}_2(x, y) \end{cases}. \tag{4.3}$$

The proof is given in Appendix C. According to the interpretation in [18], the Mumford-Shah model tries to achieve clustering based on a measure of the membership defined by $\mathcal{E}_l$, where $l = 1, 2$. Hence, the proper choice of this measure is essential to the segmentation.

4.4 Piecewise linear patch reconstruction

The following establishes the mathematical model and the associated solution for the segmentation of non-smooth image structures.

4.4.1 The model of piecewise linear patch reconstruction

In the original Mumford-Shah model, each region is modeled as a piecewise smooth surface g_l that minimizes the following functional energy:

$$g_l^* = \operatorname{argmin}_{g_l} \int_{\Omega_l} [(I - g_l)^2 + \lambda \|\nabla g_l\|] dxdy, \quad (4.4)$$

where $l = 1, 2$, λ is a penalty coefficient. A major reason for this formulation is that the energy can be minimized by solving standard partial differential equation, i.e. the diffusion equation. Besides, the energy is convex. Hence, if the solution exists it is also unique. The problem is that this formulation imposes the image smoothness.

To cope with non-smooth image structures, the image patches are considered, in this work, as linear combination of bases of patches, such as in many image appearance models. The corresponding formulation is as follows:

$$\mathbf{p}(x, y) = \sum_{k=1}^K \alpha_k(x, y) \mathbf{v}_k = \sum_{k=1}^K \langle \mathbf{p}, \mathbf{v}_k \rangle_{\boxtimes} \mathbf{v}_k, \quad (4.5)$$

with the varying weights $\alpha_k = \langle \mathbf{p}, \mathbf{v}_k \rangle_{\boxtimes}$ for different patches over the image, where $\{\mathbf{v}_k, k = 1, 2, \dots, K\}$ is a set of K orthonormal bases that reconstruct the patch, $\langle f, g \rangle_{\boxtimes} = \int_{\boxtimes} f(x)g(x)dx$ is the inner product defined over a patch region. The symbol $\boxtimes$ is used to denote the operations, such as norms and inner products, on patches. It is obvious that the exact reconstruction is possible regardless of the smoothness of the image patch. To find the bases, the reconstruction is formulated as energy minimization as follows:

$$\{\mathbf{v}_k^l\}^* = \operatorname{argmin}_{\{\mathbf{v}_k^l\}} \int_{\Omega_l} \left\| \mathbf{p} - \sum_{k=1}^K \langle \mathbf{p}, \mathbf{v}_k^l \rangle_{\boxtimes} \mathbf{v}_k^l \right\|_{\boxtimes}^2 dxdy, \quad (4.6)$$

where $\|f\|_{\boxtimes} = \left(\int_{\boxtimes} |f(x)|^p dx\right)^{\frac{1}{p}}$ is a patch norm.

Replacing the image reconstruction error in the simplified two-phase Mumford-Shah functional (4.2) with the patch reconstruction error formulated above, the formulation of piecewise linear patch reconstruction can be written as follows:

$$\min_{\{H, \{\mathbf{v}_k^1\}, \{\mathbf{v}_k^2\}\}} = \int_{\Omega} \mathcal{E}_1^{\boxtimes} H + \int_{\Omega} \mathcal{E}_2^{\boxtimes} (1 - H) dx dy, \quad (4.7)$$

where

$$\mathcal{E}_l^{\boxtimes} = \frac{1}{\|\boxtimes\|} \left\| \mathbf{p} - \sum_{k=1}^K \langle \mathbf{p}, \mathbf{v}_k^l \rangle_{\boxtimes} \mathbf{v}_k^l \right\|_{\boxtimes}^2 \quad (4.8)$$

for $l = 1, 2$, where $\|\boxtimes\|$ is the size of the patch.

From the assignment rule in Eq. (4.3), it is known that the patches will be assigned to a region if the set of bases of this region produces the patch reconstruction error smaller than that of the other regions.

4.4.2 The error of segmentation by piecewise linear patch reconstruction

In the previous subsection, it is shown that the piecewise linear patch reconstruction is capable of determining the assignment of the patches according to the patch reconstruction error regardless of the smoothness of the patch, but it cannot tell whether the assignment is correct. In the following, it is shown that the error of the assignment, i.e. the segmentation error, is bounded by the patch reconstruction error for a fixed number of, say K , bases in each segmented region. Thus, by minimizing the error of reconstruction, the error of segmentation can be reduced.

Our probabilistic analysis is based on the assumption that the patches are identically independently distributed over the regions. The assignment rule of Eq. (4.3) enables us to assess probabilistically the correctness of the segmentation. We

ask whether the segmentation by Eq. (4.3) is consistent with the ground truth. Specifically, we wish to know if $\mathbf{p}' = \mathbf{p}(x', y') \in \{\mathbf{p} | [x, y]^T \in \Omega_1\} = \mathbf{P}_1$, whether $\mathcal{E}_1(x', y') < \mathcal{E}_2(x', y')$; and if $\mathbf{p}' = \mathbf{p}(x', y') \in \{\mathbf{p} | [x, y]^T \in \Omega_2\} = \mathbf{P}_2$, whether $\mathcal{E}_2(x', y') < \mathcal{E}_1(x', y')$, where $\Omega = \Omega_1 \cup \Omega_2$ is the true partition. This requires us to analyze the following segmentation error rate:

$$\begin{aligned} \varepsilon_{seg} = & \frac{1}{|\Omega_1|} \sum_{[x', y']^T \in \Omega_1} \mathbf{1}_{[\mathcal{E}_1(x', y') > \mathcal{E}_2(x', y')]} \\ & + \frac{1}{|\Omega_2|} \sum_{[x', y']^T \in \Omega_2} \mathbf{1}_{[\mathcal{E}_2(x', y') > \mathcal{E}_1(x', y')]}, \end{aligned} \quad (4.9)$$

where $|\Omega_1|$, $|\Omega_2|$ are the sizes of the sets Ω_1 and Ω_2 .

The segmentation error rate shown above can be represented by probability, i.e. $\varepsilon_{seg} = P_{\Omega_1}[\mathcal{E}_1(x', y') > \mathcal{E}_2(x', y')] + P_{\Omega_2}[\mathcal{E}_2(x', y') > \mathcal{E}_1(x', y')]$ for sufficiently large population of Ω_1 and Ω_2 , where P_{Ω_1} is the probability of certain events due to $\mathbf{P}_1$, and P_{Ω_2} is the probability of the events due to $\mathbf{P}_2$. We expect the error rate to be small. Accordingly, the error bound stated in the following proposition for segmentation can be obtained for the proposed piecewise patch reconstruction model of Eq. (4.7).

Proposition 4.5 *If the subregions Ω_1, Ω_2 are sufficiently large, such that $E_{\mathbf{P}_1}[\mathcal{E}_1^\boxtimes] = \frac{1}{|\Omega_1|} \int_{\Omega} \mathcal{E}_1^\boxtimes H dx dy = \frac{1}{|\Omega_1|} \sum_{\Omega_1} \mathcal{E}_1^\boxtimes$, $E_{\mathbf{P}_2}[\mathcal{E}_2^\boxtimes] = \frac{1}{|\Omega_2|} \int_{\Omega} \mathcal{E}_2^\boxtimes (1 - H) dx dy = \frac{1}{|\Omega_2|} \sum_{\Omega_2} \mathcal{E}_2^\boxtimes$, then the following error bound for classifying patch $\mathbf{p}$ holds:*

$$\varepsilon_{seg} \leq \sum_{l=1,2} \frac{\|\boxtimes\| \int_{\Omega} \mathcal{E}_l^\boxtimes(x, y) H_l dx dy}{|\Omega_l| (\|\mathbf{p}\|_{\boxtimes} - R_{3-l}^2 K^{2-2q_{3-l}})}, \quad (4.10)$$

where $\|\cdot\|_{\boxtimes}$ is the patch norm, R_1, R_2, q_1, q_2 are constants, K is the number of bases and $0 \leq q_1 < \infty, 0 \leq q_2 < \infty$.

The proof is deferred to Appendix C. The above error bound guarantees in theory that the segmentation error rate due to the assignment rule of Eq. (4.3) can be

reduced by minimizing the reconstruction error with respect to a fixed number of bases.

4.4.3 Global optimal linear patch reconstruction

It has been shown that the segmentation error rate is upper bounded by the total patch reconstruction error for a fixed K . The minimization of the patch reconstruction error can therefore minimize the segmentation error. In what follows, the optimal linear patch reconstruction by minimizing the reconstruction error is to be addressed. An algorithm that leads to the global optimal solution to the problem is to be presented.

The patch reconstruction can be formulated as follows:

$$\begin{aligned} \{\mathbf{v}_k^{l*}\} &= \underset{\{\mathbf{v}_k^l\}}{\operatorname{argmin}} \int_{\Omega} \mathcal{E}_l^{\boxtimes}(x, y, \{\mathbf{v}_i^l\}) H_l dx dy \\ \text{s.t. } \forall i \neq j, \langle \mathbf{v}_i^l, \mathbf{v}_j^l \rangle_{\boxtimes} &= 0, \forall k, \|\mathbf{v}_k^l\|_{\boxtimes} = 1, \end{aligned} \quad (4.11)$$

where $l = \{1, 2\}$, $\mathcal{E}_l^{\boxtimes}(x, y, \{\mathbf{v}_k^l\}) = \|\mathbf{p} - \sum_{k=1}^K \langle \mathbf{p}, \mathbf{v}_k^l \rangle_{\boxtimes} \mathbf{v}_k^l\|_{\boxtimes}^2$, $\langle \mathbf{p}, \mathbf{v}_k^l \rangle_{\boxtimes} = \iint_{-m/2}^{m/2} \mathbf{p} \cdot \mathbf{v}_k^l dudv$ and $\|\mathbf{f}\|_{\boxtimes}^2 = \iint_{-m/2}^{m/2} \mathbf{f}^2(u, v) dudv$.

A useful identity regarding this formulation is the following:

$$\begin{aligned} &\underset{\{\mathbf{v}_k^l\}}{\operatorname{argmin}} \int_{\Omega} \mathcal{E}_l^{\boxtimes}(x, y, \{\mathbf{v}_k^l\}) H_l dx dy \\ &= \underset{\{\mathbf{v}_k^l\}}{\operatorname{argmax}} U(\{\mathbf{v}_k^l\}), \end{aligned} \quad (4.12)$$

where $U(\{\mathbf{v}_k^l\}) = \int_{\Omega} \sum_{k=1}^K \langle \mathbf{p}^l, \mathbf{v}_k^l \rangle_{\boxtimes}^2 H_l dx dy$, and $l = \{1, 2\}$, $\{\mathbf{v}_k^l\}$ are orthonormal.

Before solving the reconstruction (4.11) (or (4.12) equivalently), it is worth mentioning that *the energy minimization problem defined in Eq. (4.11) is non-convex*. Such a problem is known to have local optimal solutions [62]. However,

it is shown in the following that the global optimal solution to the linear patch reconstruction problem is attainable. The key result regarding the optimality is the following theorem. The formal statements with their proofs are deferred to Appendix C.

Although the minimization problem is constrained and nonconvex, the following shows that the global optimal solution to the problem is still attainable.

Theorem 4.6 *Given a set of functions $\{\mathbf{w}_k, k = 1, 2, \dots, K\}$ defined as follows:*

$$\begin{cases} \mathbf{w}_k = \sum_{h=1}^N \alpha_{kh} \mathbf{e}_h, \\ \forall i, j, \langle \mathbf{w}_i, \mathbf{w}_j \rangle_{\boxtimes} = 0, \text{ and } \forall i, \|\mathbf{w}_i\|_{\boxtimes} = 1, \end{cases} \quad (4.13)$$

where $\{\mathbf{e}_h, h = 1, 2, \dots, N\}$ are all the eigenvectors of Λ_{HI} , and Λ_{HI} is defined by the following:

$$\begin{aligned} \Lambda_{HI}(u, v, u', v') \\ = \int I(x + u, y + v) H_l I(x + u', y + v') H_l dx dy, \end{aligned} \quad (4.14)$$

then the following bound is true:

$$U(\{\mathbf{w}_k\}) \leq \sum_{k=1}^K \lambda_k = U(\{\mathbf{e}_k\}), \quad (4.15)$$

where $U(\{\mathbf{w}_k\})$ is defined in Eq. (4.12), and $\{\lambda_k | 1 \leq k \leq K\}$ are the first K eigenvalues of Λ_{HI} .

The proof is deferred to Appendix C. The above suggests the matrix eigen-decomposition as the proposed solution to the global optimal reconstruction. However, the eigen-decomposition typically requires converting the image to patches, computing the covariance matrix followed by Singular Value Decomposition (SVD).

In the proposed segmentation framework, the optimal reconstruction is to be performed iteratively during the segmentation. Thus, the eigen-decomposition of the matrix obtained from the extracted and labeled patches can be time-storage consuming. Alternatively, the gradient descent could be the solution. The gradient descent equation to minimize the error of Eq. (4.11) is the following:

$$\frac{\partial \mathbf{v}_n^l(u, v, t)}{\partial t} = \int_{\Omega} \left\langle I_H, \mathbf{v}_n^l \right\rangle_{\boxtimes} I_H(x + u, y + v) dx dy. \quad (4.16)$$

The reasons for this choice are the following.

Theorem 4.7 *Given the following eigenvalue problem,*

$$[\Lambda_{HI}] \mathbf{e}_k = \lambda_k \mathbf{e}_k, \quad (4.17)$$

where λ_k is the eigenvalue, $\mathbf{e}_k$ is the eigenvector and $[\Lambda_{HI}] \mathbf{x} = \int_{\{u,v\}} \Lambda_{HI} \mathbf{x} du dv$, and if there are finitely many, say N , eigenvectors of $[\Lambda_{HI}]$, the following **gradient descent procedure** for solving (4.11) (or (4.12) equivalently) converges to the global optimal solution of (4.11), with the initial bases $\mathbf{v}_0 = \sum_{k=1}^K a_k \mathbf{e}_k$ for any $\{a_k | a_k > 0, 1 \leq k \leq K\}$.

$$\begin{aligned} & \textbf{while } 1 \leq n \leq K \textbf{ do} \\ & \quad \text{STEP } n: \left\{ \begin{array}{l} \text{Solve } \mathbf{v}_n^l \text{ by (4.16)} \\ \text{s.t. } \mathbf{v}_n^l = \mathbf{v}_n^l - \sum_{k=1}^{n-1} \langle \mathbf{v}_n^l, \mathbf{v}_k^{l*} \rangle_{\boxtimes} \mathbf{v}_k^{l*}, \\ \|\mathbf{v}_n^l\|_{\boxtimes} = 1 \end{array} \right. \quad (4.18) \\ & \textbf{end while} \end{aligned}$$

The proof is deferred to Appendix C. Briefly, the above theorem says that the gradient descent method produces the eigen-patches. Moreover, the convergence of the gradient descent is linear.

Corollary 4.2 *If there are $N < \infty$ eigenvectors of Λ_{HI} , and the eigenvalues are ordered such that $\lambda_1 > \lambda_2 > \dots > \lambda_N$, the gradient descent iteration in each step of the greedy procedure defined by Eq. (4.18) converges linearly in the worst case.*

The proof is deferred to Appendix C. Besides, the gradient descent does not require computing the matrix Λ_{HI} or the matrix eigen-decomposition but only the convolutions.

4.4.4 Coupling with piecewise smooth model

It is easy to find two patches, which can be reconstructed equally well (yielding the same residue) by the same set of bases, but having significant difference in intensity. Hence, the proposed linear patch reconstruction model should not be used for differentiating the image patches that are different only in their illumination and color but similar in their local structure. To cope with both the cases via the same model, it is proposed to combine the model of linear patch reconstruction with the original piecewise-smooth Mumford-Shah model to form an integrated functional model of segmentation as follows:

$$\begin{aligned} \min_H \int_{\Omega} \mathcal{E}_1^{s \oplus \not\equiv} H + \mathcal{E}_2^{s \oplus \not\equiv} (1 - H) dx dy \\ + \nu \int_{\Omega} \|\nabla H\| dx dy, \end{aligned} \tag{4.19}$$

where $\mathcal{E}_l^{s \oplus \not\equiv} = \alpha(I - g_l)^2 + (1 - \alpha)\mathcal{E}_l^{\boxtimes}$, for $l = 1, 2$ and $0 \leq \alpha \leq 1$ is a predefined weight. The last term penalizes the complexity of the segmentation.

The numerical solution of H to the energy minimization model can be derived from the level set method, in which a signed distance function ϕ is used to generate

H through the Heaviside function, i.e. $H = H(\phi)$:

$$\begin{aligned} \frac{\partial \phi}{\partial t} = & -(\mathcal{E}_1^{s \oplus \not{s}} - \mathcal{E}_2^{s \oplus \not{s}}) \delta(\phi) \\ & + \nu \operatorname{div} \left(\frac{\nabla \phi}{\|\nabla \phi\|} \right) \delta(\phi). \end{aligned} \tag{4.20}$$

This equation implies that the reconstruction errors $\mathcal{E}_1^{s \oplus \not{s}}$ and $\mathcal{E}_2^{s \oplus \not{s}}$ are fixed when implementing curve evolution. The energy functional in (4.19) is minimized by alternately updating the image partition and updating the reconstruction error. This alternating energy minimization is commonly adopted for solving the Mumford-Shah models [18, 24, 28, 61, 63]. The convergence of such alternating energy minimization is often observed. The proof of such convergence is available [64].

4.5 Experiments

The proposed energy minimization method for segmentation can be implemented by alternating an algorithm of image partitioning and the patch reconstruction via either the SVD or the gradient descent procedure in (4.18). The curve evolution governed by Eq. (4.20) is adopted as the algorithm of image partitioning in the implementation. All the methods in the experiments are based on the curve evolution for fair comparison. For piecewise smooth image, a small penalty coefficient should be adopted for penalizing the contour length, for non-smooth image a large penalty of the contour length should be adopted for coping with the randomness.

Natural textures are examples of the non-smooth structured visual patterns. Evidences show that the texture patches can be reconstructed (modeled or represented) well by patch subspaces, i.e. the eigenfilters, as reported in [65] where the texture classification was investigated in depth. Therefore, the proposed model of piecewise linear patch reconstruction is evaluated for segmentation on a subset

Figure 4.2: (a) is the template for texture mosaicing. (b) is the initial contour for curve evolutions.

of Brodatz textures [66]. The textures chosen are the relatively *regular* textures with *similarly sized* texture stimuli (textons). Therefore, a collection of textures is obtained: {D3, D5-6, D15-22, D24, D34-36, D49, D52-53, D55, D57, D65, D68, D76-77, D79, D81-85, D101-106}, which is called the set **S**. Afterwards, two sets of 1260 mosaic texture images are generated by all possible pairings of the textures from **S** for evaluation. The template for paring the textures is shown in Figure 4.2(a). This template is also used as the ground truth for evaluating the segmentation. One of the mosaic sets is composed of the original textures, the other is the set of textures with the mean intensity subtracted. Both the sets are challenging, especially the second one.

The coupled model for segmentation of smooth contents is also evaluated, and the one-against-all strategy is adopted to cope with multiple regions of different textures.

4.5.1 Evaluation of the gradient descent patch reconstruction

In this subsection, the proposed gradient descent method for computing the optimal bases is evaluated. The objective is to compare the reconstruction errors due to the bases produced by gradient descent with the reconstruction errors due to

the bases produced by SVD.

The gradient descent method for linear patch reconstruction is evaluated based on the Set **S** and the Yale face database. The reconstruction errors are compared according to the first $k(k = 1, 2, \dots, 20)$ optimal bases produced by the gradient descent method with the errors according to the bases produced by SVD. The reconstruction errors due to SVD and the gradient descent solution can be computed by evaluating (4.11) for the entire image domain. The comparison of the reconstructions by the Gabor filter and the SVD solution is also presented. The principle of maximum filtering response is applied to select the first 20 Gabor filters from a filter bank of 40 filters. The Gabor filter bank contains 40 Gabor filters of 40 orientations uniformly distributed in $[0, \pi)$. The σ defining the spread is fixed to be 1. The patch size is 13 by 13 for Gabor filters and patches. Afterwards, the selected filters are then orthogonalized to compute the reconstruction errors. The results are shown in Figure 4.3. The results by the gradient descent are very close to those of SVD, while the results of Gabor filters do not match those of SVD at all. The initial bases for the gradient descent are randomly generated. The results of reconstruction show that without a proper choice of the scale of Gabor filters, the reconstruction by Gabor filters is unsatisfactory. The selection of the scale is known to be difficult.

4.5.2 Evaluation of the segmentation by linear patch reconstruction

This subsection presents the experiments for evaluating the segmentation by both SVD and gradient descent (GD) linear patch reconstruction. The Set **S** is used in the experiments. The proposed methods, i.e. the SVD based and GD based methods, are compared with the segmentation by PC [24] and PS model [61] [28], the Region-Scalable Fitting (RSF), a.k.a. the Local Binary Fitting (LBF) [41],

Figure 4.3: The scatter plots of the *normalized* reconstruction errors for each face or texture image by the first 1-20 bases. The value 1 corresponds to the worst reconstruction, and 0 corresponds to the perfect reconstruction. The left two plots are the orthogonalized Gabor bases (vertical) vs. SVD (horizontal); The other two are the Gradient descent (vertical) vs. SVD (horizontal). Results by gradient descent are almost as the same as SVD for different numbers of bases.

the Gabor filtering based method [55] [39] and the local histogram based Chan-Vese model [67] which is denoted as HistPC henceforth. The Gabor filtering based method could be viewed as a baseline approach for texture segmentation, and the local histogram-based method is the state-of-the-art approach. For the proposed method, 8 bases for each region are used. For the Gabor filtering based method, 8 Gabor filters are selected from a bank of 24 filters for each region according to the criterion of maximum filtering response since this criterion appears like model fitting [68]. The Gabor filter bank contains 24 Gabor filters of 8 orientations uniformly distributed in $[0, \pi)$, and the σ defining the spread is allowed to be 1, 2 and 3. The patch size is 13 by 13 for Gabor filters and the patches.

A common initial curve is used for the curve evolutions. The initial curve in the image domain is shown in Figure 4.2(b). Note that the initial contour crosses the true boundary of the two regions, and the converged contour is expected to outline the region of texture B on the right.

Figure 4.4 visualizes the comparisons in terms of the segmentation error. The proposed methods give significantly lower error rates. The quantities are also summarized in Table 4.1. It may also be observed that the segmentation errors by the SVD and the GD are comparable. The computational time for the segmentation

Table 4.1: Comparison of segmentation error summarized from Figure 4.4.

Data	Error of proposed methods		Error of other methods			
	SVD	GD	LBF	PS	HistPC	Gabor
Original	12.56±10.10	11.66±11.15	42.78 ±10.30	40.00±10.52	9.96±15.85	33.38±12.94
No mean	14.76±10.44	13.33±11.82	45.64 ±12.87	43.81±13.50	24.12±21.56	42.68±8.02

based on SVD is 0.145 ± 0.048 seconds per iteration for the two datasets, while the computational time for the segmentation based on GD is 0.137 ± 0.006 seconds, which suggests that the GD based method is preferred for segmentation compared with the SVD based method. The mean error by the local histogram based Chan-Vese active contour for the original textures is small, but the variation of its performance is large. Besides, this local histogram based method is still ineffective for differentiating the different textures having the same mean intensity. We can also observe from the results that Gabor features can deal with textures when the textures differ in their mean intensities. However, when there is little difference in the mean intensities of the textures, the Gabor features are also ineffective. The quantitative results are summarized in Table 4.1. Besides, the curve evolutions, as well as the corresponding converged patch bases, are visualized for two textured images and a picture from the Berkeley dataset [69] in Figure 4.5. The picture is composed of different structures, and the piecewise linear patch reconstruction is capable of differentiating the non-smooth structure from the piecewise smooth structure. We may observe that the optimal bases are latent region descriptors. The segmentation by other models are shown in Fig. 4.6.

4.5.3 Segmentation by the coupled model with parameter tuning

To cope with both smooth and non-smooth contents, the coupled model proposed previously should be used. The quality of the result of the segmentation by the proposed coupled model depends on the value of the parameter α . It is obvious

Figure 4.4: Comparison of segmentation error rates on the images with the mean intensity subtracted (left), and on the images of original textures (right). Proposed methods, i.e. the SVD & GD, are robust and more accurate.

Figure 4.5: Curve evolution on the Brodatz D104-22 pair (top row) and D85-106 pair (mid row) and a real image taken from Berkeley dataset with the corresponding converged patch bases. The curves are drawn in red (better to view in color). The leftmost and rightmost two columns of the patch bases correspond to the background/foreground regions. The D22, D106 and the lady are regarded as the foregrounds.

Figure 4.6: The converged curve evolution results of different methods with the same initialization

that the coupled model becomes the conventional Mumford-Shah model if $\alpha = 1$, and the model becomes the pure piecewise linear patch reconstruction if $\alpha = 0$. The two different models are applicable to different situations. α affects the final results and its effect is investigated here to find a value that produces overall good segmentations. It is proposed to select a value of α that leads to the minimum error of segmentation on both non-smooth and smooth structures.

In the experiments, the segmentation with different α values is run on all the mosaic images that are composed of the original textures (without subtracting the mean off), which is used in the previous subsection. The value of α is changed from 0.1 to 0.9 to obtain the plot of the error of segmentation shown in Figure 4.7. The range of α that produces steady performance is from 0 to 0.1, and the performance starts to drop from 0.1 onwards. At this point, it is interesting to know the performance of the coupled model on piecewise smooth image when $\alpha = 0.1$. The model with $\alpha = 0.1$ is therefore applied to a piecewise smooth image with added noise, and good result is shown in Figure 4.8, which indicates that $\alpha = 0.1$ is already sufficient for segmentation of piecewise smooth images. These suggest $\alpha = 0.1$ for both smooth and non-smooth image contents. For example, for the vessel images shown in Figure 4.9, satisfactory results of segmentation can be seen. For each of the images in Figures 4.8 and 4.9, the pairs of sets of bases are similar. We may further appreciate the advantage of using the small weight $\alpha = 0.1$ by comparing the results in Figures 4.8 and 4.9 with the results obtained by the linear

patch reconstruction alone, i.e. $\alpha = 0$. The results are shown in Figure 4.10.

Figure 4.7: The box plot of segmentation error against α

Figure 4.8: The segmentation of synthetic piecewise smooth image under noise. The images on the left show the initialization, the images in the middle show the converged curves. In the images on the right, the left two columns are the bases of the background region, and the rest correspond to the vessel region.

Figure 4.9: The segmentation of piecewise smooth images. The images on the left show the initialization, the images in the middle show the converged curves. In the image mosaics on the right, the left first two columns of mosaics from the left are the bases of the background region, and the rest correspond to the vessel region.

Figure 4.10: The segmentation of piecewise smooth images by purely linear patch reconstruction. The top row shows the results of segmentation. The bottom row shows the optimal bases corresponding to the images above. In the image mosaics of the bases, the left two columns are the bases of the background region, and the rest correspond to the foreground region (the regions enclosed by the contour curves).

4.5.4 One-against-all segmentation

The two-phase Mumford-Shah model is capable of partitioning the image into multiple regions of two types of contents. However, the two-phase model cannot cope with multiple regions of multiple types of contents. The problem of segmentation of image into multiple regions of multiple types of contents can be addressed via the one-against-all strategy. In other words, a problem of n -phase segmentation can be considered as n subproblems of two-phase segmentation. In each subproblem, the image is to be partitioned into the target region and the background region composed of all the other regions.

The one-against-all strategy for handling different groups of regions is evaluated on the mosaic images containing five different textures. The input and output of the segmentation are shown in Figure 4.11. The initial rectangular contours are shown in a unique color assigned to a region in one image. The converged contour curves are shown in the same set of colors in the other image. The optimal bases corresponding to different regions are also visualized. The resultant optimal bases well capture the principal structure of the corresponding regions. The error rates

Figure 4.11: Multi-region segmentation via the one-against-all strategy (better to view in color). The top row shows the two input-output pairs of mosaic images composed of multiple texture regions. The input is the initial rectangular contour laying over the image. In each region, a unique color is assigned to the contour. The output is the converged curve shown in the same color as in the input. The bottom rows show the optimal bases corresponding to each region in the sequence of top, bottom, left, right and center.

Table 4.2: Errors of the segmentation shown in Figure 4.11

	Top	Bottom	Left	Right	Center	Total
Figure 4.11(a)	5.16%	2.19%	1.94%	4.3%	3.2%	16.8%
Figure 4.11(b)	2.86%	2.11%	3.94%	2.03%	1.92%	12.88%

of the segmentation are summarized in Table 4.2.

4.6 Discussions

The size of the patches and the number of bases are predetermined. Both parameters affect the segmentation. For example, when increasing the patch size, the number of bases needed for achieving small reconstruction error would increase due to the enlarged dimensionality of the patch. Therefore, to achieve a small error

of segmentation the number of bases needed has to be large. A larger patch size and more bases would also provide a richer description of the non-smooth structure, which may help segmentation in complex situations. With a small patch size, which means the reconstruction error can be small with few bases, the discrimination might not be clear, since the same set of bases may give similar reconstruction errors to different groups of patches in such case. However, bigger patch size with a larger number of bases may lead to larger computational cost and larger error of boundary localization. The final selection is essentially an empirically determined one that balances between computational efficiency, boundary localization accuracy and the overall segmentation accuracy.

4.7 Concluding remarks

In this chapter, a unified energy minimization framework of non-smooth image segmentation which does not require a separate feature extraction process is proposed. Satisfactory segmentation is achieved by energy minimization of the error of piecewise linear patch reconstruction. The segmentation error rate of the proposed model is proven to be bounded by the patch reconstruction error. The gradient descent method for solving the linear patch reconstruction is proven to be globally optimal under the assumption presented in Theorem 4.7. The experiments validate the proposed theoretical claims and show clearly the superior performance of the proposed methods over other related methods.

Chapter 5

Rigid and Nonrigid Shape Evolutions for Shape Registration and Recovery

This chapter presents a novel deformable shape model based on curve evolution, which is called the Prior Variation Shape Evolution (PVSE). The PVSE can approximate general rigid and nonrigid shape deformations. The energy-minimizing PVSE equation is derived to minimize the active contour energies for locating the object boundaries that enclose object regions in images. A theory of shape preservability of the PVSE has been established for modeling nonrigid shape deformations. Due to the shape preservability, the energy-minimizing PVSE achieves shape registration for the simultaneous object segmentation and shape recovery in situations of missing parts on the shape as well as overlapping of multiple shapes. Experimental results validate the capability of the PVSE for modeling rigid and nonrigid shape deformations, and they also show the effectiveness of the simultaneous object segmentation and shape recovery by the PVSE.

5.1 Introduction

The original active contour formulations of object segmentation [3, 18] did not incorporate deformable shape modeling. In the more recent literature on application of object segmentation, such as in medical image analysis [32] and tracking [34] etc., it has been observed that the object segmentation cannot provide shape of the object of interest in the presence of occlusion and cluttered background. The occlusion and cluttered background in the scene may cause missing parts on the objects of interest. The resultant segmentation can be similar to Figures 5.1 (d), (e), and (f). In another scenario, the silhouettes of multiple objects can overlap and the observed silhouettes can be similar to those in 5.2 (d), (e), and (f). We often desire the shape silhouettes of (a), (b), and (c) of Figures 5.1 and 5.2 to be recovered simultaneously from the segmentation.

The objective of shape recovery with object segmentation may be achieved by aligning a shape representation known *a priori* to the object of interest in the image, which is referred to as a process of *shape registration*. Figure 5.3 illustrates such a process in which the deformed horse overlapped with the rider is registered and recovered according to the prototype shape(s).

Figure 5.1: (a) A toy monkey; (b) toy monkey after a rigid transformation; (c) toy monkey after a nonrigid transformation; (d), (e) and (f) are the toy monkeys in (a) (b) and (c) with missing parts.

Figure 5.2: (a) A horse; (b) The rigidly transformed horse; (c) the nonrigidly deformed horse; (d), (e) and (f) are the horses in (a), (b) and (c) merged with the rider.

Figure 5.3: A horse registration process

5.1.1 Related works

Shape registration requires a model of object boundary/region and a deformable shape model. The active contour can be considered as a boundary/region model. There exist active contour formulations based on various criteria such as the edge detection [3], the region grouping by regional homogeneity [18] or regional dissimilarity [29] etc. The optimization for active contour models is generally based on the curve evolution. Recently, there are attempts to reformulate the active contours as graph cut [25, 70]. The deformable shape model, also known as the shape prior, has been incorporated in the active contours for the shape recovery in object segmentation, such as [31–35, 71, 72]. The active shape model [73] may be viewed as one early work on the problem. However, this active shape model was not developed to cope with occlusion and cluttered background.

Among the pioneering works on shape prior based active contours, the shape

variations were conventionally modeled *statistically* in [32–35, 72]. Leventon *et al.* [32] modeled shape variation by using Principal Component Analysis (PCA), and measured the shape dissimilarity by the shape distance from a mean shape. More recently, Etyngier *et al.* [35] formulated the space of prior shapes for curve evolution as a shape manifold to which the well established nonlinear dimensionality reduction techniques are applicable. However, the statistical prior shape models require large dataset of training shapes, and every single reliable sample silhouette shape is to be delineated manually, which is labor intensive. These may restrict the application of the statistical prior shape model based active contours.

There exist active contours based on the *deterministic model* of shape variations. Chen *et al.* [31] formulated a model of simultaneous object segmentation and shape recovery as the Geodesic Active Contour with a prior term of scaling-, rotation- and translation-invariant shape dissimilarity. Cremers *et al.* [33, 34] developed the framework based on deterministic shape alignment and statistical models. However, only a simple point transformation of the linear type, i.e. the similarity transform, was considered in these works. Besides, the shape prior and the boundary location are considered as tradeoffs in these models. Thus, these models required certain parameters known beforehand, and the optimal solution could be neither the object boundary nor the desired shape. The global optimization of the active contour with translation-invariant shape prior was also considered in [36]. In contrast to all these works, the present work addresses the deterministic deformable shape models that capture general shape deformations while precluding undesirable shape variations, and it is proposed to minimize the active contour energy via such general deformable shape models.

Shape warping by warping mapping is a general deformable shape model [74–76], while similarity and affine transformations are special cases. In this thesis, the warping transformation and warping mapping are used interchangeably when

there is no risk of confusion. Warping mapping has been used for modeling shape deformations in the framework of *deformable template matching* for object boundary location [75]. Compared with the aforementioned statistical models, the shape warping has two advantages:

- the warping mapping can naturally model both the rigid and nonrigid deformations; and
- the number of reference shape samples needed can be as few as one, and the reference shape is called a prototype.

However, there are also some problems with the conventional frameworks of deformable template matching, which might have limited the application of the deterministic modeling of shape variation. For example, the measure of the optimality of the deformable template matching is formulated based on local image features on the boundary, such as edges [75, 77] and Gabor features [78], which can be restrictive comparing with the general active contour models. Besides, random optimization schemes are often used to escape from the many local optimal solutions. Consequently, the computations for a single image can be impractical, and the convergence is difficult to be determined. A deterministic algorithm for the shape registration may be preferred.

5.1.2 Contributions

In this work, a novel curve evolution model of shape deformation to minimize active contour energy is proposed for shape registration. Besides, the shape preservability of the curve evolution model is also explored.

- *Prior Variation Shape Evolution* - The shape deformation is considered as a sequence of short-time shape warpings, i.e. the shape in the next iteration is generated by *slightly* warping the present shape. This leads to a generic

mathematical model of shape deformation which is called the Infinitesimal Warping Iteration (IWI). The formulation of the IWI is general, and it is shown that the IWI can represent the similarity and affine transformations. The curve evolution equation of general IWI is then derived. The curve evolution equation is called the *Prior Variation Shape Evolution* (PVSE) equation. The vector fields of the warping mappings in IWI are named the *prior variations*.

- *A theory of shape preservability* - The question why the affine and similarity transformations can *preserve* shape are investigated. Based on the PVSE equations, it is shown that the dynamics of the zero-crossings of the curvature of the shape boundary depends only on the second-order derivatives of the prior variations. It is also shown that the second order derivatives of the prior variations of affine and similarity transformations are zero. Hence, the zero-crossing of the curvature, as a shape characteristic, *never* changes during the PVSEs of affine and similarity transformations. This completes a theory of shape preservability. Accordingly, it is argued that the PVSE is shape-preserving if the second-order derivatives of the prior variations are small. Based on these, a type of prior variations called the *prior vibrations* of bounded second-order derivatives is proposed for modeling the nonrigid shape deformations via PVSE.
- *Calculus of prior variations* - By substituting the PVSE equations into the calculus of variations, the energy minimizing PVSE equations for shape registration can be obtained. The derivation is named the calculus of prior variations. The implementation of the PVSE can be based on the well-established level set method.

There are some existing works on deterministic modeling of shape deformation in active contour [79, 80]. They proposed the parametric active contours constrained by parametric Euclidean affine or projective transformations [79, 80] respectively. The present work is distinguished from them in the following aspects:

- the novel technique of calculus of prior variations for deriving the energy minimizing PVSE equations developed in this work is compatible with any known active contour energy;
- the PVSE can be conveniently implemented by the level set method without the need for reparametrization of the contour curves;
- the PVSE models a wide class of shape deformations of both rigid and non-rigid types; and
- a mathematical theory towards the shape preservability is established to guide the modeling of shape deformation by PVSE.

5.1.3 Organizations

The rest of this chapter is organized as follows. In Section 5.2, the PVSE equation is derived from the Infinitesimal Warping Iteration, and the rationale of the PVSE for modeling shape deformations is discussed. In Section 5.3, a theory of shape preservability by PVSE is presented. Based on this theory, the PVSE with prior vibration is proposed for modeling nonrigid shape deformations. In Section 5.4, the derivation of the energy-minimizing PVSE is obtained, and the derivation of the energy-minimizing PVSE is applied to two prototypical active contours, i.e. the geodesic active contour and the Chan-Vese active contour. In Section 5.5, the PVSE for modeling shape deformation is validated by using shape context matching [81], and the PVSE for shape registration is evaluated. Both qualitative and quantitative results are reported. This chapter is concluded in Section 5.7.

5.2 Modeling shape variations by curve evolutions

In this section, the curve evolution model based on warping is proposed.

5.2.1 Infinitesimal Warping Iteration

The mathematical definition of the warping transformation is given before the proposed curve evolution model of shape warping is presented.

Definition 5.1 (Warping transformation) The warping transformation is a parametric function mapping $\tilde{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \theta)$, where $\mathbf{x} \in \mathbb{R}^2$, $\theta \in \mathbb{R}^n$ is the parameter vector, and $\mathbf{f} : \mathbb{R}^2 \mapsto \mathbb{R}^2$, such that the *identity* property, i.e. $\mathbf{x} = \mathbf{f}(\mathbf{x}, \mathbf{0})$, holds.

The identity property is necessary for the warping transformation due to the following *consistency* property according to the mean value theorem, if $\mathbf{f}$ is differentiable:

$$\begin{aligned} \|\mathbf{f}(\mathbf{x}, \theta) - \mathbf{x}\| &= \|\mathbf{f}(\mathbf{x}, \theta) - \mathbf{f}(\mathbf{x}, \mathbf{0})\| \\ &\leq \|\nabla \mathbf{f}(\mathbf{x}, \alpha\theta + (1 - \alpha)\mathbf{0})\| \|\theta\| \\ &\leq \max_{0 \leq \alpha \leq 1} (\|\nabla \mathbf{f}(\mathbf{x}, \alpha\theta)\|) \|\theta\|, \end{aligned} \quad (5.1)$$

which ensures that small parameters yield small shape deformations.

The shape deformation is considered as a dynamic process of shape warpings as follows, which would be common in computer vision tasks such as tracking:

$$\mathbf{x}(p, t^{k+1}) = \mathbf{f}(\mathbf{x}(p, t^k), \theta(t^k + \tau)), \quad \|\theta(t^k + \tau)\| \leq \epsilon, \quad (5.2)$$

where $\mathbf{x}(p, t^k) = C(p, t^k)$ is the point on the shape contour, $\mathbf{f}(\cdot, \cdot)$ is a parametric warping transformation, $\theta(\cdot)$ is the vector of warping parameters and $\epsilon \rightarrow 0$. Eq. (5.2) is named the *Infinitesimal Warping Iteration* (IWI) due to the small θ . To

follow the identity property of warping, $\theta(t^k) = 0$ is required. In this model, the contour shape is slightly deformed at each iteration of warping due to the small θ .

This general deformable shape model can represent a large variety of shape variations according to the relatively arbitrary form of the warping transformation $\mathbf{f}$. It will also be shown that the fundamental rationale of this model in (5.2) is its capability of modeling the geometrical transformations, such as the similarity and affine transformations. Let us consider the following general formulation of Infinitesimal Linear Warping Iteration:

$$\begin{aligned}\mathbf{x}(p, t^{k+1}) &= \mathbf{A}^{k+1} \mathbf{x}(p, t^k) + \mathbf{b}^{k+1}, \\ \mathbf{x}(p, t^0) &= \mathbf{x}_o(p).\end{aligned}\tag{5.3}$$

Based on the formulation above, the following facts show the fundamental rationale of the IWI model.

Proposition 5.6 (Generalized closure property) *For any k ,*

$$\mathbf{x}(p, t^k) = \tilde{\mathbf{A}} \mathbf{x}_o(p) + \tilde{\mathbf{b}},\tag{5.4}$$

in which $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{b}}$ as follows are constants for any p :

$$\tilde{\mathbf{A}} = \prod_{j=1}^k \mathbf{A}^j, \text{ and, } \tilde{\mathbf{b}} = \sum_{i=1}^k \prod_{j=i}^k \mathbf{A}^j \mathbf{b}^i.\tag{5.5}$$

Corollary 5.3 (Generalized closure of affine transformation) *If for all j , $\mathbf{A}^j$ is invertible, then $\tilde{\mathbf{A}}$ is also invertible.*

The affine transformation, as a parametric warping mapping, can be defined as follows:

$$\begin{aligned}\tilde{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \mathbf{b} \\ &= \begin{bmatrix} 1 + a_{11} & a_{12} \\ a_{21} & 1 + a_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} a \\ b \end{bmatrix},\end{aligned}\tag{5.6}$$

where $\mathbf{A}$ is an invertible matrix, and $\mathbf{b}$ is a translation vector. As a special case of the above, the closure property of the similarity transformation is the following.

Proposition 5.7 (Generalized closure of similarity transformation) *If $\mathbf{x}_1 = e^{\lambda_1} \mathbf{R}(\omega_1) \mathbf{x}_0 + \mathbf{b}_1$, and $\mathbf{x}_2 = e^{\lambda_2} \mathbf{R}(\omega_2) \mathbf{x}_1 + \mathbf{b}_2$, then we have $\mathbf{x}_2 = e^{\lambda_3} \mathbf{R}(\omega_3) \mathbf{x}_0 + \mathbf{b}_3$, where $\lambda_3 = \lambda_1 + \lambda_2$, $\omega_3 = \omega_1 + \omega_2$, $\mathbf{b}_3 = \mathbf{b}_2 + e^{\lambda_2} \mathbf{R}(\omega_2) \mathbf{b}_1$, $\mathbf{R}(\omega)$ is a rotation matrix function with the parameter of orientation ω .*

Specifically, the similarity transformation is the following:

$$\tilde{\mathbf{x}} = \begin{bmatrix} e^\lambda \cos(\omega) & \sin(\omega) \\ -\sin(\omega) & e^\lambda \cos(\omega) \end{bmatrix} \mathbf{x} + \begin{bmatrix} a \\ b \end{bmatrix},\tag{5.7}$$

for any $\mathbf{x}$, where $\theta_{rg} = [\lambda, \alpha, a, b]^T$, and the origin of the transformation is placed at the centroid of the evolving shape, in which the identity property of warping holds.

5.2.2 Prior variation shape evolution

In the IWI, the small displacement due to small θ allows the Taylor expansion of the warping for $\tau \rightarrow 0$ at each iteration to be as follows:

$$\begin{aligned}\mathbf{x}(p, t^{k+1}) &= \mathbf{x}(p, t^k) + \left[\frac{\mathbf{Df}(\mathbf{x}(p, t^k), \theta)}{\mathbf{D}\theta} \right] \bigg|_{\theta=0} \theta(t^k + \tau).\end{aligned}\tag{5.8}$$

The above motivates us to approximate the IWI by curve evolution. Let $t^{k+1} = t^k + \tau$, and taking derivative w.r.t. τ at $\tau = 0$ on the two sides of the above, the differential form of the IWI model can be written as follows:

$$\left. \frac{\partial \mathbf{x}(p, t^k + \tau)}{\partial \tau} \right|_{\tau=0} = \left[\frac{\mathbf{Df}(\mathbf{x}(p, t^k), \theta)}{\mathbf{D}\theta} \right] \bigg|_{\theta=\mathbf{0}} \left. \frac{d\theta}{d\tau} \right|_{\tau=0}. \quad (5.9)$$

Let $t = t^k + \tau$, we have $\frac{dt}{d\tau} = 1$. Hence, the following curve evolution equation of the IWI can be obtained by replacing $\mathbf{x}(p, t)$ with $C(p, t)$ as follows:

$$\left. \frac{\partial C(p, t)}{\partial t} \right|_{t=t^k} = [\mathbf{V}^\theta] \left. \frac{d\theta}{dt} \right|_{t=t^k}, \quad (5.10)$$

and

$$[\mathbf{V}^\theta] = \left[\frac{\mathbf{Df}(C(p, t), \theta)}{\mathbf{D}\theta} \right] \bigg|_{\theta=\mathbf{0}}, \quad (5.11)$$

where $[\mathbf{V}^\theta]$ emphasizes that $\mathbf{V}^\theta$ is a $2 \times n$ matrix-valued function, and n is the dimension of θ . Note that $[\mathbf{V}^\theta]$ depends only on the present contour location. Hence, the form of it can be chosen *a priori*. Eq. (5.10) is the *Prior Variation Shape Evolution* (PVSE) equation. The PVSE equation can be implemented via the level set method to avoid the artificial reparametrization. The level set equation corresponding to Eq. (5.10) is the following:

$$\begin{aligned} \frac{\partial \phi}{\partial t} &= -\nabla \phi \frac{\partial C(p, t)}{\partial t} = \left\langle \nabla \phi, [\mathbf{V}^\theta] \left. \frac{d\theta}{dt} \right|_{t=t^k} \right\rangle, \\ \phi(x, y, 0) &= \phi_o(x, y), \end{aligned} \quad (5.12)$$

where $\|\nabla \phi(x, y, t)\| = 1$, ϕ is the signed distance function.

To understand the effect of PVSE more intuitively, the curve evolution of similarity transformation and affine transformation are visualized in Figures 5.4 and 5.5 respectively.

Figure 5.4: The curve evolution for concurrently rotation, scaling and translation of “A”.

Figure 5.5: The curve evolution for affine transformation of “A”.

5.3 A theory of shape preservability

In this section, the shape preservability of the PVSE is investigated in order for the shape recovery. The formulation of PVSE is based on the given prior variations. The prior variations are vector fields of warping transformations. The general formulation of a warping transformation often consists of a rigid part $\mathbf{f}_{rg}$ and a nonrigid part $\mathbf{f}_{nrg}$ with transformation parameters $\theta = [\theta_{rg}^T, \theta_{nrg}^T]^T$, where θ_{rg} and θ_{nrg} are both column vectors. A review of many popular warping mappings is in [76]. The warping mapping considered in this chapter is the *parametric warping transformation* [76].

The rigid part of a warping transformation models the visual perception of the motion of rigid shape. The rigid part is not necessarily a distance-preserving transformation, i.e. the rigid transformation in the strictest sense. In the following, the rigid part will be referred to as the *rigid transformation* for conciseness if there is no risk of confusion. Every point on a rigid shape is only allowed to move with the same transformation and the same parameters. According to the choice of geometry, one may formulate the rigid part as the similarity transformation, affine transformation or projective transformation.

The nonrigid part is called the *nonrigid deformation* which models the visual

perception of the nonrigid deformation of a nonrigid shape. To complete the proposed curve evolution based deformable shape model, we only need to choose the form of the IWI for the rigid and nonrigid parts. Choosing the form of IWI is equivalent to choosing the prior variations for the PVSE. The prior variations of the rigid part can be chosen as the prior variations of either the similarity transformation in (5.6) or the affine transformation in (5.7). The formulations of the corresponding prior variations is to be given later. We have yet to choose the prior variations for nonrigid deformations that is shape-preserving. In the following, a theory for choosing the nonrigid prior variations is established.

5.3.1 The invariance of shape characteristics

Before choosing the nonrigid prior variations that preserve shapes, the shape preservability of shape deformations is to be investigated. It is proposed to quantify the invariance of shape characteristics under PVSEs, which forms a mathematical model of shape preservability. It is to be shown that the mathematical model of shape preservability is valid for the PVSEs of similarity and affine transformation.

An important shape characteristic is the zero-crossings of curvature in the curvature scale space [82, 83]. It is of interest to investigate how such zero-crossings, or feature points, move under the geometrical transformations. Particularly, the invariance of the characteristics by the geometrical transformations in the form of PVSEs is to be examined. A fundamental result is the following.

Theorem 5.8 *Given the curve evolution governed by (5.10), the dynamic of the curvature, κ , at the positions where $\kappa = 0$, i.e. the zero-crossings of the curvature, satisfies the following:*

$$\left\| \frac{\partial \kappa}{\partial t} \right\|_{\kappa=0} \leq \left\| \frac{d\theta}{dt} \right\| \sum_{i=1}^n \left\| \frac{\mathbf{D}^2 \mathbf{V}^{\theta_i}}{\mathbf{D} \mathbf{x}^2} \right\|. \quad (5.13)$$

In words, the dynamics of the curvature of the zero-crossing points is bounded by the norm of the second-order derivatives of the prior variations of the PVSEs. Besides, it can also be shown that the distance between any two contour parameters of the zero-crossings is also bounded by the norm of the second-order derivative of the prior variations.

Corollary 5.4 *Given two zero-crossing points $C(p_1, t)$ and $C(p_2, t)$ at time t , the dynamic of the distance between the parameters, p_1 and p_2 , of the feature points is bounded as follows:*

$$\left| \frac{d}{dt} |p_1 - p_2| \right| \leq \frac{2}{\|\kappa_s C_p\|} \left\| \frac{d\theta}{dt} \right\| \sum_{i=1}^n \left\| \frac{\mathbf{D}^2 \mathbf{V}^{\theta_i}}{\mathbf{D} \mathbf{x}^2} \right\|. \quad (5.14)$$

The proofs of Theorem 5.8 and Corollary 5.4 are deferred to Appendix D. The theoretical findings above suggest *the use of PVSEs with prior variations of small second order derivatives for shape recovery.*

In the following, it is verified that the similarity and affine transformations can preserve the zero-crossings of the shape curvature. The prior variations of the similarity PVSE, which is referred to as the *similarity variations*, are the following:

$$\mathbf{V}^\lambda = [x, y]^T, \mathbf{V}^\omega = [y, -x]^T, \quad (5.15)$$

$$\mathbf{V}^a = [1, 0]^T, \mathbf{V}^b = [0, 1]^T. \quad (5.16)$$

The prior variations of the PVSE of affine transformation, which is referred to as the *affine variations*, are the following:

$$\mathbf{V}^A(x) = \left[[x, 0]^T, [y, 0]^T, [0, x]^T, [0, y]^T \right]^T, \quad (5.17)$$

$$\mathbf{V}^b(x) = \left[[1, 0]^T, [0, 1]^T \right]^T. \quad (5.18)$$

It is obvious that the the second order derivatives of these prior variations vanish,

meaning that the zero-crossings of the curvature does not change. Note that the similarity and affine transformations preserves shape visually. Hence, the established theory is validated for similarity and affine transformations.

5.3.2 The prior vibrations

From the previous discussions, it is proposed to choose the PVSE with prior variations of small second order derivatives to model the nonrigid deformations. The following family of functions are considered.

$$\begin{bmatrix} e_{mn}^1(\mathbf{x}) \\ e_{mn}^2(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} \frac{\sin(\pi nx) \cos(\pi my)}{\pi^2(n^2+m^2)} \\ \frac{\cos(\pi mx) \sin(\pi ny)}{\pi^2(n^2+m^2)} \end{bmatrix}. \quad (5.19)$$

This class of functions has been adopted in the deformable template matching [74, 75]. Variants of this function class are also used, such as in [84, 85]. The chosen prior variations are in the form of these functions as follows:

$$[\mathbf{V}_{nrg}^\theta] = [\mathbf{V}_{mn} | 0 < m \leq M, 0 < n \leq N], \quad (5.20)$$

$$\mathbf{V}_{mn} = \begin{bmatrix} e_{mn}^1(C(p, t)) \\ e_{mn}^2(C(p, t)) \end{bmatrix}^T, \quad (5.21)$$

which is termed the *prior vibrations* due to the periodic oscillating structure of the basis functions.

To validate this choice of non-rigid prior variations, the following theorem is used.

Theorem 5.9 *For the function class defined in (5.20) and (5.21), the following*

bound of its second order derivatives holds:

$$\sum_{m,n=1}^{MN} \left\| \frac{\mathbf{D}^2 \mathbf{V}_{mn}}{\mathbf{D} \mathbf{x}^2} \right\|_{\infty} \leq 4MN, \quad (5.22)$$

where M and N are defined in (5.20), and $\|\cdot\|_{\infty}$ is the ∞ -norm of matrix.

The proof of Theorem 5.9 is deferred to Appendix D. According to the established theory of shape preservability presented in Section 5.3.1, one may choose small values of M, N , i.e. using the truncated series, to preserve the shape characteristic. Likewise, the smoothness of the prior variations, in terms of the magnitude of the first order differential of $\mathbf{V}_{mn}$, can also be guaranteed by using small values of M, N , which is consistent with the theory presented in [86].

Accordingly, the PVSE equation with the prior vibrations is as follows:

$$\begin{aligned} \frac{\partial C}{\partial t} = & \sum_{m,n=1}^{M,N} \begin{bmatrix} e_{mn}^1(C(p,t)) \\ 0 \end{bmatrix}^T \begin{bmatrix} \frac{d\theta_{mn}^1(t)}{dt} \\ \frac{d\theta_{mn}^2(t)}{dt} \end{bmatrix} \\ & + \begin{bmatrix} 0 \\ e_{mn}^2(C(p,t)) \end{bmatrix}^T \begin{bmatrix} \frac{d\theta_{mn}^1(t)}{dt} \\ \frac{d\theta_{mn}^2(t)}{dt} \end{bmatrix}, \end{aligned} \quad (5.23)$$

where the coordinates are normalized such that $\mathbf{x} = [x, y]^T \in [0, 1] \times [0, 1]$ as in [74].

5.4 Calculus of prior variations

This section addresses the problem: how can the PVSE be used to minimize active contour energies based on the calculus of variations? Since the prior variations can be chosen to make the PVSE preserve shapes and object boundary can be located by minimizing the active contour energies, the resultant energy minimizing PVSE should be able to achieve simultaneous object segmentation and shape recovery.

The calculus of variations, which can be represented by the *Gâteaux* derivative, is as follows:

$$\frac{d\mathcal{J}}{dt} = \left\langle \nabla \mathcal{J}, \frac{\partial C}{\partial t} \right\rangle_p^*, \quad (5.24)$$

where $\nabla \mathcal{J}$ is the functional gradient, $\langle \cdot, \cdot \rangle_p^*$ is the functional inner product of two 2-D vector functions.

By substituting the PVSE equation (5.10) into the *Gâteaux* derivative in Eq. (5.25), the following is obtained:

$$\begin{aligned} \frac{d\mathcal{J}}{dt} &= \left\langle \nabla \mathcal{J}, \left[\mathbf{V}^\theta \right] \frac{d\theta}{dt} \right\rangle_p^* \\ &= \sum_{i=1}^n \left\langle \nabla \mathcal{J}, \mathbf{V}^{\theta_i} \right\rangle_p^* \frac{d\theta_i}{dt} = \frac{\mathbf{D}J}{\mathbf{D}\theta} \frac{d\theta}{dt}, \end{aligned} \quad (5.25)$$

where $\theta = [\theta_1, \theta_2, \dots, \theta_n]^T$, n is the number of parameters and $t = t^k$ is omitted.

The gradient descent equation for θ is as follows:

$$\begin{aligned} \frac{d\theta}{dt} &= -\frac{\mathbf{D}J}{\mathbf{D}\theta} = -\int_C \left[\mathbf{V}^\theta \right] \nabla \mathcal{J} dp \\ &= - \left[\left\langle \mathbf{V}^{\theta_i}, \nabla \mathcal{J} \right\rangle_p^* \middle| i = 1, 2, \dots, n \right]^T. \end{aligned} \quad (5.26)$$

By back substituting (5.26) into the PVSE equation of (5.10), one obtains the following.

$$\begin{aligned} \frac{\partial C(p, t)}{\partial t} &= - \sum_{i=1}^n \left\langle \nabla \mathcal{J}, \mathbf{V}^{\theta_i} \right\rangle_p^* \mathbf{V}^{\theta_i} \\ C(p, 0) &= C_o(p). \end{aligned} \quad (5.27)$$

The equation of the energy minimizing PVSE (5.27) finalizes the proposed approach for shape registration, i.e. the simultaneous object segmentation and

shape recovery. The functional gradients of various active contour energies can be substituted into (5.27) and the corresponding level set equation (5.12). For example, the gradient descent equation of GAC [6, 17] is the following:

$$\frac{\partial C}{\partial t} = -\nabla \mathcal{J}_{GAC}(C) = g\kappa \mathbf{N} - \langle \nabla g, \mathbf{N} \rangle \mathbf{N}. \quad (5.28)$$

The general form of the corresponding PVSE equation is the following:

$$\frac{\partial C}{\partial t} = -\sum_{\theta_i \in \theta} \langle \nabla \mathcal{J}_{GAC}, \mathbf{V}^{\theta_i} \rangle_p^* \mathbf{V}^{\theta_i}. \quad (5.29)$$

Likewise, we may rewrite the gradient descent equation of the Chan-Vese active contour [24] or the region competition [18] as follows:

$$\begin{aligned} \frac{\partial C}{\partial t} &= -\nabla \mathcal{J}_{CV}(C) \\ &= (\mu\kappa - \nu + \lambda_2|u_0 - c_2|^2 - \lambda_1|u_0 - c_1|^2) \mathbf{N}, \end{aligned} \quad (5.30)$$

where $\mu, \nu, \lambda_1, \lambda_2$ are positive penalty coefficients, u_0 is the (feature) image, and c_1 and c_2 are defined below:

$$c_1 = \frac{\int_{\Omega} u_0 H(\phi) dx dy}{\int_{\Omega} H(\phi) dx dy}, \quad c_2 = \frac{\int_{\Omega} u_0 (1 - H(\phi)) dx dy}{\int_{\Omega} (1 - H(\phi)) dx dy}. \quad (5.31)$$

The general form of the corresponding PVSE equation is the following:

$$\frac{\partial C}{\partial t} = -\sum_{\theta_i \in \theta} \langle \nabla \mathcal{J}_{CV}, \mathbf{V}^{\theta_i} \rangle_p^* \mathbf{V}^{\theta_i}. \quad (5.32)$$

It can be observed from the equation above that the velocity of the energy-minimizing curve evolution is generated by *projecting* the negative functional gradient to the prior variations. The derivations for the energy-minimizing PVSE

Algorithm 2 The algorithm of energy-minimizing PVSE

Input: C_o , $[\mathbf{V}_{rg}^\theta]$ or $[\mathbf{V}_{nrg}^\theta]$, $g(\cdot) : \mathcal{D} \mapsto (0, \mathbb{R}+)$ or $I(\cdot) : \mathcal{D} \mapsto [0, 255]$
Output: ϕ^*

- 1: $\phi^0(x, y) \leftarrow \begin{cases} 0, & [x, y]^T \in \{C_o = \partial\Omega_o\} \cap \mathcal{D} \\ -1, & [x, y]^T \in \Omega_o \cap \mathcal{D} \\ 1, & [x, y]^T \in \overline{\Omega_o} \cap \mathcal{D} \end{cases}$
- 2: $[\mathbf{V}^\theta] \leftarrow [\mathbf{V}_{rg}^\theta]$ or $[\mathbf{V}_{nrg}^\theta]$
- 3: $k \leftarrow 1$
- 4: **While** NOT CONVERGENCE **do**
- 5: $\phi^k \leftarrow \begin{cases} \|\nabla \phi^k\| = 1 \\ \phi^k \circ (\phi^{k-1})^{-1}(0) = 0 \end{cases}$
- 6: $\nabla \mathcal{J}^k \leftarrow \nabla \mathcal{J}_{GAC}(\phi^k, g)$ or $\nabla \mathcal{J}_{CV}(\phi^k, I)$
- 7: $\mathcal{P}_{\mathbf{V}^\theta}^k \leftarrow \sum_{\theta_i \in \theta} \langle \mathbf{V}^{\theta_i}, \nabla \mathcal{J}^k \rangle_p^* \mathbf{V}^{\theta_i}$
- 8: $\phi^{k+1} \leftarrow \phi^k - \Delta t \langle \mathcal{P}_{\mathbf{V}^\theta}^k, \nabla \phi^k \rangle \delta_\epsilon(\phi^k)$
- 9: $k \leftarrow k + 1$
- 10: **End While**
- 11: $\phi^* \leftarrow \phi^k$

---

equations for all the known active contours are similar. The previously obtained rigid and nonrigid prior variations (5.15), (5.17) and (5.19) can be substituted into (5.29) and (5.32) for implementation. The energy-minimizing PVSE equation is to be implemented by numerical solvers, such as the level set method as in Eq. (5.12). The convergence can be detected if the evolving curve has little motion for a few iterations. The implementation is summarized in Algorithm 2.

5.5 Experimental results

5.5.1 Evaluation of the PVSE equation for modeling shape deformations

In previous sections, the PVSE equation is derived from the IWI. The rationale of the PVSE equation for being a deformable model is discussed theoretically. In this

section, the experimental validation of the PVSE for modeling shape deformations is presented.

5.5.1.1 Experimental settings

The objectives of the experiments are summarized as follows:

- it is of interest to know whether the PVSE with prior similarity variations (5.15) and prior affine variations (5.17) can represent the similarity and affine transformations; and
- it is of interest to know the capability of the PVSE with prior vibrations for modeling non-rigid deformations.

To achieve the above two objectives, it is proposed to measure the shape distance between the evolving shape and the prototype shape. The shape distance ρ is defined as follows:

$$\rho(\mathbf{S}, \mathbf{S}_o) = \min_{\mathbf{A}, \mathbf{b}} d(\mathbf{S}, \mathbf{A}\mathbf{S}_o + \mathbf{b}), \quad (5.33)$$

where d is the Hausdorff distance, and $\mathbf{A}$ and $\mathbf{b}$ can be the parameters of similarity or affine transformation, which specify the geometry of the shape space. In 5.33, the natural distance between two point sets, i.e. the Hausdorff distance, can be adopted to compute the distance d . The Hausdorff distance may be sensitive to noise, but the shapes are clean in the context of this work. It is also not required to consider the effect of noise since the deformation of a prototype shape cannot be noisy. The transformation parameters of $\mathbf{A}, \mathbf{b}$ are determined by *shape context* matching [81].

In principle, if the PVSEs with prior variations (5.15) and (5.17) can achieve the corresponding shape transformations, the shape distance according to the corresponding similarity or affine geometry will be small. Moreover, the non-affinity, in terms of the affine shape distance, should be large for non-rigid deformation.

A baseline result is also needed. In the experiments, the bolded **A** in Figure 5.6 (left) is used as the prototype, and the slanted *A* (actually in freestyle) in Figure 5.6 (middle) is used as the baseline shape. The size of both images is 64×64 . The optimally aligned shape against the reference shape is also shown in Figure 5.6(right). The affine invariant Hausdorff shape distance between the optimally aligned shape and the reference shape is about 4.5 by (5.33).

Figure 5.6: The prototype (left) and baseline (middle) shapes used for evaluation, and the right shows the affine shape context matching of the baseline (blue crosses) by the prototype (red circles).

The experiment proceeds as follows:

- (a) *generate random velocities $\frac{d\theta}{dt}$ for the prior variations (both the rigid type and the non-rigid type);*
- (b) *run PVSE for a prototype shape and record the evolving shapes;*
- (c) *perform the shape context matching for the recorded shapes and estimate the similarity/affine transformation parameters; and*
- (d) *compute the shape distances defined by (5.33).*

In detail, 10 random vectors of $\frac{d\theta}{dt}$ are generated for the PVSE with the prior variations in (5.15) and (5.17). The prior vibrations in (5.19) are called the 1st order prior vibrations if $M + N = 1$, and likewise the rest are the 2nd, 3rd, ..., 10th order prior vibrations. Ten random vectors for each of the 10 prior vibrations are generated. Each of the PVSEs are run for 50 iterations to generate a collection of $(10 + 2) \times 10 \times 50 = 6000$ shapes for evaluation.

Figure 5.7: The shape distance for the 50 iterations of similarity (left) and affine (right) PVSEs. The red line is the shape distance between the two “A”s in Figure 5.6.

5.5.1.2 Experimental results and analysis

The quantitative results of the experiments are shown in Figures 5.7 and 5.8. The results are visualized by boxplots for all iterations. The top and bottom of each box are the 25th and 75th percentiles of the distribution respectively. The line in the middle of each box is the median, and the red crosses are outliers. By inspecting Figure 5.7, it can be observed that the shape distances are small compared with the baseline value and the curve evolution does not increase the shape distance, which is good. To conclude, the PVSE is effective for simulating the similarity and affine transformations. It may also be concluded that the PVSE with prior vibrations is non-rigid according to Figure 5.8. The non-affinity is relatively stable for different order of prior vibrations as shown in 5.8(left). As the curve evolution iterates, the non-rigidity, in terms of non-affinity, increases according to Figure 5.8(right). Some examples of the simulated PVSEs are visualized in Figure 5.9 for better understanding the experiments. Table 5.1 summarizes the quantitative results corresponding to Figure 5.9.

Figure 5.8: The shape context distance for the 50 iterations of the Prior Vibration Curve Evolutions of order 1 to order 10. The left shows the shape distances for each order of prior vibration; The right shows the shape distances at each iteration for all orders. The red line is the shape distance between the two “A”s in Figure 5.6.

Table 5.1: Hausdorff shape distance corresponding to Figure 5.9

Rows	Distance type	Distances		
		(a)	(b)	(c)
First	Similarity	1.85	1.36	1.18
Second	Affine	2.44	1.54	1.66
Third	Affine	1.73	3.87	5.71

Figure 5.9: The PVSE and the affine shape matching by shape context. The blue curves are the contour curves due to PVSE, the red crosses are the results of shape matching by shape context. The first row corresponds to a PVSE by similarity variations; The second row corresponds to a PVSE by affine variations; The last row corresponds to a PVSE by 3rd order prior vibrations.

5.5.2 Evaluation of the PVSE for shape registration

5.5.2.1 Experimental settings

In this section, the proposed algorithm for shape registration, i.e. the simultaneous object segmentation and shape recovery, is evaluated. An experiment to validate the presented theoretical formulations quantitatively is conducted. The experiments are designed with the following objectives in mind:

- the PVSE is to be evaluated under the situation that there are no missing parts or merge shapes;
- the PVSE is to be evaluated in the presence of missing parts and/or merging shapes; and
- different rigid and non-rigid deformations should be considered.

The size of the images for the monkey and the monkey with missing parts,

Table 5.2: The distributions of randomly generated parameters from uniform distributions

Transformations	Parameters	Distributions
Rigid, $\vec{x} \in [-1, 1] \times [-1, 1]$	θ	$U(-\pi/4, \pi/4)$
	λ	$U(-0.9, 1.1)$
	a, b	$U(-0.1, 0.1)$
Non-rigid, $\vec{x} \in [0, 1] \times [0, 1]$	ξ_{mn}^k , $1 \leq m + n \leq 3$, $k = 1, 2$	$U(-0.125, 0.125)$

the horse and the horse merged with the rider as shown in Figure 5.1 and Figure 5.2, is 128×128 . The parameters are randomly generated for 50 similarity transformations and 50 random non-rigid warpings in the form of (5.19). Then the corresponding rigid and non-rigid transformations are applied to the monkey and horse shapes to obtain a total of $200 + 200 = 400$ shapes for evaluation. The probability density distributions that are used to generate random parameters are summarized in Table 5.2. In the experiment, the Chan-Vese active contour is chosen as the model of segmentation.

5.5.2.2 Experimental results

The results of the experiments are shown in Figures 5.10 and 5.11, in which the abbreviations are summarized in Table 5.3. The statistical evaluations based on the Jaccard's coefficients are given by the box-plots in Figure 5.11. The registration performance as measured by the Jaccard's coefficients indicate the similarity between registration output and the deformed ground truths. Generally the registration performance based on the shape similarity measure is high. From Figures 5.10 and 5.11, it can also be seen that the differences between the initial prototype shapes and the shapes to be registered do not show any correlated effect on the performance of registration. This shows the insensitivity of the proposed algorithm to such differences.

For a better understanding of the proposed algorithm, some examples of the

Table 5.3: Abbreviations

MR:	rigidly transformed ground truth of monkey
MRM:	rigidly transformed monkey with missing parts
MNR:	nonrigidly transformed ground truth of monkey
MNRM:	nonrigidly transformed monkey with missing parts
HR:	rigidly transformed ground truth of horse
HRM:	rigidly transformed horse merged with rider
HNR:	nonrigidly transformed ground truth of horse
HNRM:	nonrigidly transformed horse merged with rider

curve evolution are presented in Figures 5.12, 5.13 and 5.16. Figures 5.12 and 5.13 are two sample results from the experiment of Section 5.5.2.1. It can also be noted that the Jaccard coefficient for the TPAMI logo in Figure 5.16 is 0.81.

The results of object shape extraction from a picture of tattoo in Figure 5.15 and a picture of a clownfish in Figure 5.17(a) are also presented. For the clownfish, the PVSE is applied to the red channel of this picture, which is shown in Figure 5.17(b). In the implementation, the affine PVSE is used followed by a 3rd order Prior Vibration Shape Evolution. The shape evolution process in Figure 5.18. The first row shows the affine PVSE, the second row shows the non-rigid PVSE during which the local shape deformation at the tail can be observed. There is no quantitative results for the tattoo and clownfish, as the ground truth shapes are not available. However, the results are visually neat.

Figure 5.10: Shape Variation: the distances between the initial prototype shapes and the shapes to be registered. This is to show that the deformations of shapes is relatively large. The task is challenging

Figure 5.11: Registration performance measured by Jaccard coefficients (vertical axis).

Figure 5.12: PVSE with prior vibrations for registration of the toy monkey. (a) to (e) show the curve evolution of 90 iterations until convergence at (e).

5.6 Discussions on the limitations

In this work, it is assumed that the object detection has been achieved manually or automatically. Based on the detection, one may obtain an initial guess of the object shape which initialized the shape registration. The rigid and nonrigid shape registration is achieved with only one available sample of the object shape. To represent the object shape of interest, some more samples may be needed to either construct the shape model or to evaluate the shape model.

Figure 5.13: PVSE with prior vibrations for registration of the horse shape. (a) to (e) show the curve evolution of 150 iterations until convergence at (e).

Figure 5.14: PVSE with prior vibrations for registration of another horse shape. (a) to (e) show the curve evolution of 500 iterations until convergence at (e).

Figure 5.15: Shape registration with affine PVSE to find the “h” in the tattoo on the skin. (a) to (e) show the curve evolution of 200 iterations until convergence at (e).

5.7 Concluding remarks

In this chapter, a novel deformable shape model based on curve evolution, i.e. the PVSE, is proposed. The PVSE models a wide class of shape deformations of both rigid and nonrigid types. A theory of shape preservability of PVSE is also

Figure 5.16: Segmentation and recovery of the TPAMI logo from heavy noise and occlusion by PVSE with the 6th-order prior vibration. The top-left is the input image. The bottom-left is the convergent curve laying over the ground truth. (a) to (h) visualize the curve evolution of 150 iterations until convergence at (h).

Figure 5.17: The box of the clownfish overlaying the original color image and the red channel to which the PVSE applies.

Figure 5.18: Shape registration with PVSE to find the clownfish in the coral. The first row shows the affine shape evolution until convergence. The second row shows the prior vibration shape evolution until convergence.

established. The energy-minimizing PVSE equation to minimize active contour energies is also derived by using the proposed technique of calculus of prior variations. The developed shape evolution algorithm can handle concurrent occurrence of shape deformation, missing parts and shape merges. Experiments show promising results. The novel technique of calculus of prior variations for deriving the energy minimizing PVSE equations developed in this chapter is compatible with any known active contour energy to the best knowledge of the authors. The PVSE equations can be conveniently implemented by the level set method without the need of reparametrization of the contour curves.

Chapter 6

Matching-Constrained Active Contours

In using active contours for object segmentation, the region of interest is often required for the initial contour. This is often provided by the user. This chapter presents a novel constrained optimization model of object segmentation leading to an automated active contour framework without user intervention. The idea is to use feature point matching to estimate the boundary contour of the object of interest. The boundary estimation by feature point matching is considered as the initial feasible solution to a constrained optimization problem giving rise to a novel matching-constrained active contour model. The projected-gradient descent equations are derived to solve the constrained optimization problem exactly. Numerical experiments show that the proposed matching-constrained active contour achieves automatic object segmentation in complex real images given the feature points and the shape of the object.

6.1 Introduction

Object segmentation is formulated as a contour optimization problem in active contour models. Given the initial contours sufficiently close to the boundaries of the object of interest, the active contour can converge to the object boundaries. However, initialization by the user is not desired for a fully automatic system. Automatic object segmentation without user intervention is an ultimate objective of the research on object segmentation. This chapter presents an algorithm that can automatically perform segmentation of the object of interest without user intervention.

6.1.1 Related works

The edge-based active contour models, e.g. the Geodesic Active Contour [17], requires the initial contours provided by the user to be sufficiently close to the boundaries. The indication of the edges of interest is crucial. Region-based active contours may be insensitive to initialization, e.g. the Chan-Vese model [24]. However, several image regions may share the same regional property, and the discriminating regional property of an object of interest in the image is difficult to be modeled exactly beforehand. The difficulty in modeling may reduce the suitability of the region-based active contour for object segmentation in real images.

Relaxation of the user initialization in edge-based active contours for object segmentation has become a research topic. Xu and Prince [87] proposed the gradient vector field (GVF) to extend the localized gradients such that the large gradients at the boundary can influence the active contour far away from the edges. This allows the initial contours to be far from the edges. However, the active contour in GVF cannot find the boundary of interest when there are clutters near the object boundaries. Paragios et al. [42] also applied the GVF to the level set based

active contours to extract multiple objects. However, they reported that their method may not perform well when the initialization is arbitrary. Li et al. [45] considered the segmentation of the GVF. Xie and Mirmehdi [43] proposed a curve evolution formulation based on edge detection, which allows more flexible contour initialization, as an alternative to the conventional edge-based active contours.

The problem of restrictive curve initialization has also been addressed based on the observations on the geometry of the gradient field in Chapter 3. However, even though the initial contours can be more flexible, the input of the object of interest from the users is still required. A previous attempt [88] for automatic initialization of the edge-based active contours selects the initial contour that has approximately the minimum energy among a set of possible contours. In other words, it is assumed that all of the stronger gradients in the image are from the boundaries of the objects of interest. This assumption is valid for the images studied in [88] but it cannot be generalized.

The global optimization techniques for some region based and edge-based active contour models have been proposed in [89, 90]. The global search does not require initial contour. The idea behind these approaches is to approximate the exhaustive search of the optimal shape candidate over the entire image domain. Hence, it achieves automatic object segmentation. The shape candidate might be referred to as shape priors in the literature of active contours. The major problem behind this idea is that the coincidence of the global optimal contour and the boundary of the object of interest is by assumption. Actually, the object of interest is often one of the local optimal solution of the active contour models. For example, the global optimal solution to GAC may be a dot, but the object of interest in fact occupies a large area.

Apart from the issues on automatic object segmentation, the shape prior modeling in active contours has also become an attractive research direction. There are

some advantages of the active contours with shape priors. For example, the shape prior modeling can help recover the shape of the object of interest in the image in the presence of occlusion, and it can also regularize the originally ill-defined active contour models, such as the geodesic active contour model [90]. Both rigid and nonrigid shape variations have been dealt with in the literature of shape prior based active contours, e.g. [31–33, 35, 86]. Leventon et al. in [32] proposed the PCA based shape prior models which capture the simple shape variations from samples. Chen et al. [31] incorporated simple shape transformations, i.e. rotation, scaling and translation, in shape prior modeling. Cremers et al. [33] considered both shape transformations and statistical shape prior models. More recently, Etyngier et al. [35] models the shape prior space as a shape manifold. A novel method which uses general rigid and non-rigid shape deformations for minimizing active contour energies is also presented in Chapter 5. However, these methods are based on gradient descent and the initialization is still crucial for most of the edge-based active contour energies. For example, in Figure 6.1(b), the prior variation shape evolution [86] is stuck at the suboptimal solution, because of the local image gradients ¹.

6.1.2 Methodology

This chapter addresses the automatic object segmentation following a divide-and-conquer strategy. The problem of automatic object segmentation is considered as a unity of two subproblems: the object detection and the boundary location. Unlike the global search methodology in [89, 90], it is arguable that the object detection can provide the coarse location of the object. The segmentation of this object is expected to be obtained with the initialization provided by the detector. The philosophy is that *if the object detection cannot be done properly, then the object*

---

¹The method in [86] allows users to select their preferred deformations for minimizing arbitrary active contour energies, and the affine transformation and GAC energy are adopted there

Figure 6.1: The importance of initialization. (a) shows a bad initialization which gives an unsatisfactory result of the segmentation in (b). (c) shows a good initialization which gives an encouraging result in (d).

segmentation is also hopeless. Such principle has not been formulated in active contours. There exist many effective object detectors based on classification, e.g. the Haar-like feature based detectors [91] and the histogram of oriented gradient feature based detectors [92]. However, classifiers often require a large number of samples for training. The complexity of the classification also varies for different objects. Object matching based on locally discriminative feature matching finds a set of points on the target object of interest in the image [93–95], achieving coarse object localization. The principle of feature point matching is illustrated in Figure 6.2. The feature point matching often follows interest point detection on a pair of images as in [93]. The similarity between all possible pairs of matches are computed. Finally, the matching algorithm will assign a set of feature points in one image, possibly a subset of the entire feature set, to a set of feature points in

Figure 6.2: Illustration of the feature point matching (with errors).

Figure 6.3: Object segmentation through active contour with shape prior initialized by feature point matching.

the other image according to the point-wise similarity. The point correspondences are visualized by the blue lines in Figure 6.2, and the corresponding points are highlighted while the other interest points are neglected. A correspondence is often represented by using a binary variable, i.e. 1 means matched and 0 means unmatched.

In this work, the object-detection-by-matching approach is adopted since the number of reference objects required for matching can be small. Besides, matching can also help estimating the parameters of the transformation from the reference object to the object in image, which may lead to an estimation of the boundary location to initialize the active contours. The main idea is illustrated in the block diagram of Figure 6.3. In the proposed framework, feature point matching is used to provide an estimation of the boundary contour of the object of interest. The

boundary estimation is used for initialization of active contours. A difficulty is that the matched feature points cannot directly determine the object boundary. Besides, the initialized contour is to be optimized for object segmentation in the conventional active contour framework, meaning that the contour should be represented as a parametric curve or an implicit contour defined by an implicit function. Lastly, the entire framework is to be formulated as a single optimization problem.

6.1.3 Contributions

The contributions of this chapter mainly lie in three aspects.

- **Affine-invariant points-to-shape alignment** The reference object and the target object in the image may be different in their locations, orientations and sizes etc. Using the methods such as in [93–95] for feature point matching, as illustrated in Figure 6.2, only the correspondences of the feature points on the reference object to those on the target object are established, achieving only the alignment of the interior feature points. By assuming that the interior points and the object shape undergo the same affine transformation from the reference object to the object in the image, a mathematical model of the affine-invariant relation between the interior feature points and the boundary contour is formulated based on kernel function approximation. Upon point alignment, the interior-points-to-shape relation gives the shape alignment.
- **Affine active contour driven by interior point motion** Based on the established affine-invariant interior-points-to-shape relation, the affine active contours driven by the affine motion of interior feature points is also obtained. The formulation involves parameterized implicit surface, which makes the optimization different from the conventional geometric active contours based

on the level set method or the parametric active contours based on splines. In this chapter, the corresponding gradient descent equations in terms of the affine parameters of the implicit surface are derived for minimizing the active contour energies.

- **Matching-constrained active contour model** The affine-invariant interior-points-to-boundary-contour relation together with the affine active contour gives rise to a unified constrained optimization model in which the score of feature point matching is used to constrain the affine active contour. The initial contour produced by feature point matching provides the initial feasible solution to the constrained optimization. The projected-gradient descent equations are derived to solve the constrained optimization exactly.

6.1.4 Organization

The rest of the chapter is organized as follows. Section 6.2 introduces the proposed mathematical model of the affine-invariant interior-points-to-shape relation, which leads up to the affine points-to-shape alignment. Section 6.3 presents the unified constrained optimization model of the framework of automatic object segmentation, namely the matching-constrained active contour, and the projected gradient algorithm. In Section 6.4, the proposed method of automatic object segmentation is evaluated on real images of cluttered scenes, and the associated quantitative analysis with discussions is presented. This chapter is concluded in Section 6.5.

6.2 Modeling affine-invariant interior-points-to-shape relation

In this section, the mathematical model of the relation between interior points and the outlining contour shape is presented. This enables the contour alignment based on feature matching.

6.2.1 Points-to-shape relation as a binary classifier

The reference object shape can be represented by its silhouette. The shape silhouette of an object can be defined as a binary function $H_o(x, y)$ as follows:

$$H_o(x, y) = \begin{cases} 1, & [x, y] \in \Omega \\ 0, & [x, y] \in \overline{\Omega} \end{cases}, \quad (6.1)$$

where Ω is the object region and $\overline{\Omega}$ is the non-object region. This definition can be used to define the object boundary contour C as follows:

$$C = \{x, y \mid \|\nabla H_o(x, y)\| \neq 0\}, \quad (6.2)$$

where $H_o = H(\phi(x, y))$, H is a Heaviside function and ϕ is a signed distance function. Generally, $\|\nabla H_o(x, y)\| = \delta(\phi(x, y))$, where δ is a Dirac delta function. In other words, the implicit shape representation is employed in this work, which is motivated by the level set method [14]. The contour in the reference image is a set of discrete points, and the continuous contour shape can be estimated by function approximation as follows:

$$H_e^* = \underset{H_e \in \mathcal{S}}{\operatorname{argmin}} \mathcal{E}(H_e) \quad (6.3)$$

$$\mathcal{E}(H_e) = \int_{\mathcal{D}} |H_o - H_e|^2 dx dy,$$

where $\mathcal{D}$ is the entire image domain, H_e is the approximate of H_o and $\mathcal{S}$ denotes the solution space of H_e .

In our context, H_e is generated by the set of the position vectors of feature points $\{\vec{\mathbf{p}}\} = \{\vec{p}_1, \vec{p}_2, \dots, \vec{p}_N\}$, i.e. $H_e = H_e(\{\vec{p}\})$. If it is possible to consider the feature points $\{\vec{\mathbf{p}}\}$ as randomly distributed points, the radial basis function (RBF) then can be used for approximating the function H_o . This idea is formulated as follows:

$$H_e(\vec{z}) = H \left(\sum_{i=1}^N \alpha_i \psi(\|\vec{z} - \vec{p}_i\|) + \beta \right), \quad (6.4)$$

where $\vec{z} = [x, y]^T$, H is a Heaviside function and we may call $\phi(\vec{z}) = \sum_{i=1}^N \alpha_i \psi(\|\vec{z} - \vec{p}_i\|) + \beta$ the *shape decision function* in which $\psi(\cdot)$ is a kernel function, and $\{\alpha_i\}$ and β are the weights and bias to be determined. The points $\{\vec{\mathbf{p}}\}$ are the center points of the kernel functions. The sign of shape decision function determines whether a point belongs to the shape.

Combining the previous formulations in Eqs. (6.1), (6.3) and (6.4), a binary classification problem is obtained based on RBF neural network in which the decision function is formed by only some of the positive samples, i.e. the feature points, but the training is accomplished by using both positive and negative samples over the entire image domain. It is considered that the training is by direct minimization of the fitting error in (6.3) with respect to the parameters $\{\alpha_i\}$ and β . The gradient descent equations for learning the parameters are given in Appendix E. The detailed learning strategy is presented in Section 6.4. An example of the trained shape decision function and the associated contour curve are shown in Figure 6.4.

Figure 6.4: A 2D visualization of the trained shape decision function overlaid the corresponding implicit contour (in red). The larger the intensity of the image is the more likely the point belongs to the shape region.

6.2.2 Imposing affine invariance in the points-to-shape relation

In real images, the object shape and the positions of the interior feature points are often different from those given in the training session. Given the matched set of feature points in the image, we are left to determine the object shape. It is assumed that the contour shape of the object undergoes the same transformation as that of the interior feature points, and especially the affine transformation is considered. This property is illustrated in Figure 6.5. This assumption is common in computer vision. The behavior that the shape and points undergo the same affine transformation is the *affine invariance property*. However, it is not sure for now whether the implicit contour of the trained H_e , or the shape decision function $\phi(\cdot)$, have the affine invariance property. This issue is addressed in the following.

Figure 6.5: Illustration of the need of affine invariance. The point correspondences should lead to the shape alignment.

The Gaussian function is chosen as the RBF kernel. Thus, the shape decision function can be written as follows:

$$\phi(\vec{z}) = \sum_{i=1}^N \alpha_i \exp \left(-\frac{\|\vec{z} - \vec{p}_i\|^2}{\sigma^2} \right) + \beta. \quad (6.5)$$

The contour curves can be defined by the shape decision function according to $C = \{x, y | \phi(x, y) = 0\}$, which is equivalent to Eq. (6.2) where H_o is replaced by H_e .

Let us consider the affine transformation of the kernel centers $\vec{p}_i$ as follows:

$$\vec{q}_i = \mathbf{A}\vec{p}_i + \vec{b}, \text{ for all } i, \quad (6.6)$$

where $\mathbf{A}$ is an invertible matrix, and $\vec{b}$ is a translation vector. The corresponding shape decision function in terms of the transformed kernel centers can be written as follows:

$$\phi(\vec{z}, \{\mathbf{A}, \vec{b}\}) = \sum_{i=1}^N \alpha_i \exp \left(-\frac{\|\vec{z} - \mathbf{A}\vec{p}_i - \vec{b}\|^2}{\sigma^2} \right) + \beta. \quad (6.7)$$

The affine transformation of the contour points can be represented by $\vec{z}_s = \mathbf{A}\vec{z}_c + \vec{b}$, where $\vec{z}_c \in C$, such that $\phi(\vec{z}_c) = 0$ in (6.5). By substituting $\vec{z}_s$ into Eq.

(6.7), we obtain the following:

$$\phi(\vec{z}_s, \{\mathbf{A}, \vec{b}\}) = \sum_{i=1}^N \alpha_i \exp\left(-\frac{\|\mathbf{A}\vec{z}_c - \mathbf{A}\vec{p}_i\|^2}{\sigma^2}\right) + \beta.$$

Regarding the affine invariance property, the above leads to the following.

Proposition 6.8 *There exists infinitely many $\vec{z}_c$ and $\vec{p}_i$, such that for each i , we have $\exp\left(-\frac{\|\mathbf{A}\vec{z}_c - \mathbf{A}\vec{p}_i\|^2}{\sigma^2}\right) \neq \exp\left(-\frac{\|\vec{z}_c - \vec{p}_i\|^2}{\sigma^2}\right)$, if $\mathbf{A}$ is not an orthogonal matrix.*

The proof is deferred to the appendix.

This means that the affine transformation of the contour curve defined by the trained shape decision function, $\phi(\vec{z})$, may not be the contour curve defined by the transformed shape decision function $\phi(\vec{z}, \{\mathbf{A}, \vec{b}\})$, which is a violation of the aforementioned affine invariance property. To address this problem, a revised shape representation ϕ_S is proposed as follows:

$$\begin{aligned} \phi_S(\vec{z}, \{\mathbf{A}^{-1}, \vec{b}\}) \\ = \sum_{i=1}^N \alpha_i \exp\left(-\|\mathbf{A}^{-1}(\vec{z} - \mathbf{A}\vec{p}_i - \vec{b})\|^2\right) + \beta. \end{aligned} \tag{6.8}$$

Substituting $\vec{z}_s$ into (6.8), it can be verified that for general $\mathbf{A}$ and $\mathbf{b}$, the following affine invariance holds:

$$\phi_S(\vec{z}_s, \{\mathbf{A}^{-1}, \vec{b}\}) = \phi_S(\vec{z}_c, \{\mathbf{I}^{-1}, \vec{0}\}) = \phi(\vec{z}_c) = 0,$$

where $\mathbf{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $\vec{0}$ defines an identity transformation. ϕ_S is parameterized by $\mathbf{A}^{-1}$ rather than $\mathbf{A}$ to simplify the later derivation for optimizing $\mathbf{A}$.

We may further appreciate the significance of the affine invariance property by the example shown in Figure 6.6. Suppose we are given the pair of initial points and shape in Figure 6.4, the the interior feature points of the leaf are then

transformed by a predetermined affine transformation. The implicit contour curve defined by Eq. (6.7) is shown in Figure 6.6(a). Note that Eq. (6.7) does not have the affine invariance property. The resultant shape, which tends to be a circle, differs significantly from a shape of a leaf. With the same transformation of the feature points, the implicit contour curves defined by the affine invariant shape decision function in Eq. (6.8) is shown in Figure 6.6(b). Obviously, the affine-invariant implicit contour in Figure 6.6(b) is like a leaf transformed resulting from the same transformation applied to the feature points, which is the affine-invariance property. With such an affine invariant interior-points-to-shape relation in Eq. (6.8), we can achieve the affine points-to-shape alignment. A result of the points-to-shape alignment is shown in Figure 6.7 based on the feature matching presented in Figure 6.2.

Figure 6.6: Significance of affine invariance of the shape representation. Blue dots are the transformed feature points. The outlining curves are the implicit shape contours.

Lastly, some of the advantages of the established interior-points-to-shape model over the explicit transformation of the contour shape are enumerated as

- (a) the affine invariant interior-points-to-shape relation provides continuous implicit contours without reparametrization, which can simplify the implementation; and

Figure 6.7: A demo of points-to-shape alignment.

- (b) sometimes, the entire region of object shape, namely the shape silhouette, is of interest. However, the explicit transformation generally cannot provide a uniform region of shape. The proposed interior-points-to-shape relation provides the shape silhouette by definition.

6.3 The matching-constrained active contour

The proposed model of affine-invariant interior-points-to-shape relation can be used for contour alignment given a point matching. The implicit contours defined by the affine-invariant interior-points-to-shape relation are continuous and they do not require reparametrization when the shape changes. Such continuous geometrical shape representation is also desired by the active contour framework for segmentation. Therefore, the conventional active contour model can be reformulated by using the proposed affine invariant points-to-shape relation. Due to the affine invariance, the solution to the reformulated active contour model will be an affine transformation of the initial contour. Hence, an affine shape prior adheres to the reformulated active contour model.

In this section, a general active contour model based on the affine-invariant interior-points-to-shape relation is proposed, and a constraint of matching is used as a constraint to the model.

6.3.1 Active contour model with matching constraint

Given the interior-point-to-shape relation, the affine contour alignment by affine point alignment can be achieved. The affine point alignment can be obtained by using the matching correspondences. The aligned contour can then be used to initialize the curve evolution in the affine active contour. These form an algorithmic framework for automatic object segmentation. In this section, an optimization model is formulated to unify the matching, alignment and curve evolution in a unified optimization framework. The motivation of the proposed constrained optimization model lies in the notion of *initial feasible solution*. Specifically, the contour curve given by the feature points matching and alignment provides the initial feasible solution to the constrained optimization model of active contour. The local features matching can also be used to constrain the optimization. The model is as follows:

$$\begin{aligned}
 C^* &= \underset{C}{\operatorname{argmin}} J(C(\{\vec{\mathbf{q}}\})) \\
 \text{s.t.: } &E(\{\mathbf{q}\}) \leq \tau, \quad \vec{q}_i = \mathbf{A}\vec{p}_i + \vec{b},
 \end{aligned} \tag{6.9}$$

where $\{\vec{\mathbf{p}}\} = \{\vec{p}_1, \vec{p}_2, \dots, \vec{p}_N\}$ and $\{\vec{\mathbf{q}}\} = \{\vec{q}_1, \vec{q}_2, \dots, \vec{q}_N\}$ are the feature points on the template and the transformed sets of these feature points, $C(\{\vec{\mathbf{q}}\})$ denotes the implicit relation between the interior points and the contour shape defined by (6.8). $J(C)$ is the abstract form of the active contour energy, $E(\cdot)$ is the abstract form of the cost of joint matching and alignment, and τ is a tolerance level. Note that the set of $\{\vec{\mathbf{q}}\}$ that corresponds to the minimal value of $E(\cdot)$ must be feasible to the inequality if there exists at least one feasible solution. By this formulation, it is assumed that the optimal contour can be obtained by transforming the boundary of the template object, and it is assumed that the transformation is affine.

6.3.2 Projected gradient decent algorithm

Without knowing the actual formulations of the objective functional J and the constraint E , it is still possible to derive the solution to the abstract optimization problem based on the projected gradient descent method. The projected-gradient descent equations can be written as follows:

$$\frac{d\mathbf{A}}{dt} = \left(\nabla_{\mathbf{A}} J - \left\langle \nabla_{\mathbf{A}} J, \frac{\nabla_{\mathbf{A}} E}{\|\nabla_{\mathbf{A}} E\|} \right\rangle \frac{\nabla_{\mathbf{A}} E}{\|\nabla_{\mathbf{A}} E\|} \right), \quad (6.10)$$

$$\frac{d\vec{b}}{dt} = \left(\nabla_{\vec{b}} J - \left\langle \nabla_{\vec{b}} J, \frac{\nabla_{\vec{b}} E}{\|\nabla_{\vec{b}} E\|} \right\rangle \frac{\nabla_{\vec{b}} E}{\|\nabla_{\vec{b}} E\|} \right). \quad (6.11)$$

Besides, the following property holds:

$$\frac{dE}{dt} = 0 \quad \text{and} \quad \frac{dJ}{dt} \leq 0. \quad (6.12)$$

The derivation of this property is deferred to the Appendix E. The above indicates that the projected gradient descent algorithm governed by Eqs. (6.10) and (6.11) can reduce the value of J while leaving the value of E unchanged. This might be too strong for the inequality constrained optimization. In fact, it is only required that E in the constraint is to be smaller than a predefined tolerance τ . Therefore, only the original gradient descent is needed if $E < \tau$, and the full projected gradient algorithm is needed if $E \approx \tau$. To implement the projected gradient descent algorithm, the explicit form of the gradients $\nabla_{\mathbf{A}} J$, $\nabla_{\mathbf{A}} E$, $\nabla_{\vec{b}} J$ and $\nabla_{\vec{b}} E$ are needed. The detailed formulations are given in the following.

6.3.3 Gradient of active contour parameterized by affine motion of interior points

The projected gradient descent algorithm requires the explicit form of the gradients of the active contour energy, i.e. $\nabla_{\mathbf{A}}J$ and $\nabla_{\vec{b}}J$. The derivation of the explicit form of $\nabla_{\mathbf{A}}J$ is complex. Alternatively, we can obtain $\nabla_{\mathbf{A}}J$ numerically by using $\nabla_{\mathbf{A}^{-1}}J$. The explicit form of $\nabla_{\mathbf{A}^{-1}}J$ and $\nabla_{\vec{b}}J$ can be written as follows:

$$\nabla_{\mathbf{A}^{-1}}J = \left\langle \nabla J, -\frac{\vec{N}}{\langle \vec{N}, \nabla_{\vec{z}}\phi_S \rangle} \frac{D\phi_S}{D\mathbf{A}^{-1}} \right\rangle_p, \quad (6.13)$$

$$\nabla_{\vec{b}}J = \left\langle \nabla J, -\frac{\vec{N}}{\langle \vec{N}, \nabla_{\vec{z}}\phi_S \rangle} \frac{D\phi_S}{D\vec{b}} \right\rangle_p, \quad (6.14)$$

in which $\nabla_{\vec{z}}\phi_S$ can be computed either numerically or exactly. The closed form expressions of $\frac{D\phi_S}{D\mathbf{A}^{-1}}$ and $\frac{D\phi_S}{D\vec{b}}$ are needed for the computation, which are as follows:

$$\frac{D\phi_S}{D\mathbf{A}^{-1}} = -\sum_{i=1}^N 2w_i(\vec{z}) \frac{\vec{v}_i(\vec{z})\vec{z}^T}{\sigma^2},$$

$$\frac{D\phi_S}{D\vec{b}} = -\sum_{i=1}^N 2w_i(\vec{z}) \frac{(\mathbf{A}^{-1})^T \vec{v}_i(\vec{z})}{\sigma^2},$$

where

$$w_i(\vec{z}) = \exp\left(-\frac{\|\mathbf{A}^{-1}\vec{z} - \vec{p}_i - \mathbf{A}^{-1}\vec{b}\|^2}{\sigma^2}\right),$$

$$\vec{v}_i(\vec{z}) = \mathbf{A}^{-1}\vec{z} - \vec{p}_i - \mathbf{A}^{-1}\vec{b}.$$

The derivation is deferred to the Appendix E. The derivation is not restricted to any specific active contour model. For the Geodesic Active Contour (GAC) [17], the functional gradient, ∇J , is the following:

$$\nabla J_{GAC} = \langle \nabla g, \vec{N} \rangle \vec{N} - g\kappa \vec{N},$$

where g is an edge indicator function in which the stronger edge corresponds to smaller value, $\vec{N}$ is the normal of the contour, and κ is the contour curvature.

6.3.4 Joint formulation of point matching and alignment

To obtain the explicit form of $\nabla_{\mathbf{A}}E$ and $\nabla_{\vec{q}}E$, the explicit formulation of E in the constraint is required. E is a measure of the optimality of feature matching and point alignment. A joint optimization model of matching and alignment can be written as follows:

$$\begin{aligned} \{\vec{\mathbf{q}}^*\} = \operatorname{argmin}_{\{\vec{\mathbf{q}}\}} E(\vec{\mathbf{q}}) &= \operatorname{argmin}_{\{\vec{\mathbf{q}}\}} \sum_{ij} c_{ij} \delta_{\epsilon}(r_{ij}) \\ \text{s.t. } \forall i, \sum_j \delta_{\epsilon}(r_{ij}) &= 1, \end{aligned} \quad (6.15)$$

where δ_{ϵ} is the approximation of the Dirac delta with a parameter ϵ , $r_{ij} = \|\vec{q}_i - \vec{q}_j^t\|$ and $\{\vec{\mathbf{q}}^t\}$ is the set of target feature points in the target image, and c_{ij} is the cost of the matching between $\vec{q}_i$ and $\vec{q}_j^t$. $\delta_{\epsilon}(r_{ij})$ is the relaxed matching indicator. It has the following property:

$$\delta_{\epsilon}(r_{ij}) \approx \begin{cases} M, & r_{ij} = 0 \\ 0, & r_{ij} \neq 0 \end{cases},$$

where M is a constant. The normalization constraint on $\delta_{\epsilon}(r_{ij})$ requires all the points in the template to be matched.

In this model, the optimal δ_{ϵ} is determined by the closeness between $\vec{q}_i$ and $\vec{q}_j^t$. Thus, by aligning $\vec{q}_i$ toward a proper $\vec{q}_j^t$, the matching cost can be minimized. However, the above model is easily trapped by degenerate solutions where $\delta_{\epsilon}(r_{ij}) =$

0. To avoid such degeneration, we have the following reformulation:

$$\begin{aligned} \{\vec{\mathbf{q}}^*\} &= \operatorname{argmax}_{\{\vec{\mathbf{q}}\}} \mathcal{E}(\{\vec{\mathbf{q}}\}) = \operatorname{argmax}_{\{\vec{\mathbf{q}}\}} \sum_{ij} e^{-c_{ij}} \delta_\epsilon(r_{ij}) \\ \text{s.t. } \forall i, \sum_j \delta_\epsilon(r_{ij}) &= 1. \end{aligned} \quad (6.16)$$

The constraint in terms of $\mathcal{E}$ is therefore $E = -\mathcal{E} \leq \tau \leq 0$.

The gradients of the energy function in (6.16) are as follows:

$$\nabla_{\mathbf{A}} E(\mathbf{A}, \vec{b}) = - \sum_{ij} \hat{c}_{ij} g_{ij}(\mathbf{A}\vec{p}_i + \vec{b} - \vec{q}_j^t) \vec{p}_i^T, \quad (6.17)$$

$$\nabla_{\vec{b}} E(\mathbf{A}, \vec{b}) = - \sum_{ij} \hat{c}_{ij} g_{ij}(\mathbf{A}\vec{p}_i + \vec{b} - \vec{q}_j^t), \quad (6.18)$$

where $\hat{c}_{ij} = e^{-c_{ij}}$. The Gaussian function is adopted to approximate the Dirac delta. Thus, $g_{ij} = \delta_\epsilon(r_{ij})$. The Gaussian functions are normalized according to the constraint $\sum_i \delta_\epsilon(r_{ij}) = 1$. These two gradients can also be used for local refinement of $\mathbf{A}$ and $\vec{b}$ after the initial shape alignment.

Based on the detailed formulations presented so far, the implementation of the matching-constrained active contour is described in Algorithm 3. The initial $\mathbf{A}_o$ and $\vec{b}_o$ are computed by using established point correspondences. The locally affine matching [95] is used to provide the initial point correspondences. The matching score due to the locally affine matching is guaranteed to be high enough and the locally affine matching is robust to outliers.

6.4 Experiments

In the implementation of the matching constrained active contour, the Scale Invariant Feature Transform (SIFT) interest point detector and the SIFT feature [93] are used. The point correspondence is produced by the method reported in [95].

Algorithm 3 Projected gradient active contour

Input: Input image I , source points $\{\vec{p}^s\}$, initial $\mathbf{A}_o, \vec{b}_o$, trained $\phi_S(\cdot)$.

Output: H_e^*

```

    %Points-to-shape alignment
1: Solving  $[\mathbf{A}_o \vec{b}_o] [\mathbf{P} \vec{1}]^T = \mathbf{Q}^t$  w.r.t.  $\mathbf{A}_o, \vec{b}_o$ 
2: repeat
3:   Update  $\mathbf{A}_o, \vec{b}_o$  by Eqs. (6.17) and (6.18)
4: until Convergence
5:  $\vec{p}_i^{t_0} = \mathbf{A}_o \vec{p}_i^s + \vec{b}_o$ 
6: Evaluate  $\phi_S(\vec{z}, \{\mathbf{A}_o^{-1}, \vec{b}_o\})$  on image domain by Eq. (6.8)
    %Constrained optimization
7:  $k = 1$ 
8: repeat
9:   if  $E(\{\vec{p}_i^{t_{k-1}}\}) < \tau$  then
10:    Update  $\mathbf{A}_k$  and  $\vec{b}_k$  by Eqs. (6.13) and (6.14)
11:   else
12:    Update  $\mathbf{A}_k$  and  $\vec{b}_k$  by Eqs. (6.10) and (6.11)
13:   end if
14:   Compute  $\phi_S(\vec{z}, \{\mathbf{A}_k^{-1}, \vec{b}_k\})$  image domain by Eq. (6.8)
15:    $k = k + 1$ 
16: until Convergence

```

$\mathbf{P} = [\vec{p}_1, \vec{p}_2, \dots, \vec{p}_N]^T$ is an $N \times 2$ matrix of feature vectors in the template image, $\mathbf{Q}^t = [\vec{q}_1^t, \vec{q}_2^t, \dots, \vec{q}_N^t]^T$ is an $N \times 2$ matrix of the corresponding feature vectors in the target image.

Since the proposed method only requires the established point correspondences, the method is not confined to the choice of SIFT and the matching algorithm. The τ in (6.9) is set in relation to the matching score E from initial object matching.

In the experiments, the real images are taken from Mikolajczyk's homepage², Caltech computer vision archive³ and the ETHZ Toys dataset to evaluate the proposed method for automatic object segmentation.

6.4.1 Affine invariant shape modeling

Following the shape training process presented in Section 6.2, one can obtain the parameters of the affine invariant points-to-shape model. The template objects

---

²<http://lear.inrialpes.fr/people/mikolajczyk/>

³<http://www.vision.caltech.edu/html-files/archive.html>

Figure 6.8: Three pairs of template objects (left) and the corresponding trained shape contours overlaying the training shapes (right)

and the trained shape contours are shown in Figure 6.8. The fitting errors during the training of the RBF for the shapes are shown in Figure 6.9. The training of the shape models converges stably.

Figure 6.9: Minimization of the errors of shape modeling corresponding to Figures 6.8 (a) (b) and (c).

There is a parameter σ in the RBF based shape representation. To select the optimal σ , one may use the fitting errors with fixed initial weights α_i and threshold β for selecting σ . Specifically, the σ which corresponds to the highest fitting score ($1 - \text{normalized error}$) is selected from a set of values $\{1, 1.1, \dots, 19.9, 20\}$ containing 191 candidates. Figure 6.10(a) shows the fitting scores w.r.t. σ for the three shapes. To validate this choice, a set of 50 random values of σ from 1 to 20 is generated for an exhaustive search of the optimal σ for the leaf shape, in which the gradient descent learning is performed for each of the σ to get the convergent shape models corresponding to the 50 random values of σ . The scores of the optimal fitting, in terms of the Jaccard shape similarity, given the 50 random

Figure 6.10: Experimental results on selection of the parameter σ . (a) is the plot of fitting scores (vertical) v.s. σ (horizontal), given the fixed initial weights and thresholds. (b) is the plot of fitting scores w.r.t. different values of σ given the fixed initial weights and thresholds (blue square with curve) comparing with the fitting scores due to the optimization of the weights and threshold given the randomly selected σ (red cross with curve).

values of σ are shown in Figure 6.10(b). It can be observed that the peaks of the two curves are quite close, which means that the criterion for selecting σ values with fixed weights and threshold is effective. The major advantage of the selection of σ with fixed weights and threshold over the exhaustive search is the computational efficiency.

Upon the accomplishment of the shape model training with parameter selection, one can validate the claimed affine invariance of the shape model defined in Eq. (6.8). The trained leaf shape model is taken for evaluation. The principle is to compare the contour shape implicitly defined by the transformed interior points with the explicitly transformed contour shape, which could be considered as the ground truth shapes. The normalized Hausdorff distance, i.e. the Hausdorff distance normalized by the maximum distance between the two point sets, is chosen as the shape distance. One hundred sets of parameters for the affine transformation are generated without translation (shifting the shape decision function will not result in errors in the shape modeling), leading up to 100 randomly

generated leaves. Some examples of the implicit contours overlaying the explicitly transformed contour points are shown in Figures 6.11(a)-(i). It can be observed that the proposed interior-points-to-shape relation defined in Eq. (6.8) is affine invariant. In Figures 6.11(j)-(r), the implicit contours due to the shape model without affine invariance, i.e. Eq. (6.7), are also presented. The shape distances corresponding to Figure 6.11 are summarized in Table 6.1. In Figure 6.12, the shape distances corresponding to the two shape models for all the 100 examples are plotted. It can be observed that even when the shape transformation is large, the distance between the explicitly and implicitly transformed contour shapes is still small for the affine-invariant shape model, whereas the shape model without the affine invariance deviates significantly from the explicitly transformed shapes.

Table 6.1: Shape distances corresponding to the Figure 6.11.

Figure 6.11	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)
NHD	0.108	0.086	0.046	0.054	0.053	0.033	0.049	0.056	0.036
Figure 6.11	(j)	(k)	(l)	(m)	(n)	(o)	(p)	(q)	(r)
NHD	0.180	0.116	0.113	0.130	0.154	0.080	0.100	0.114	0.126

NHD = Normalized Hausdorff distance

6.4.2 The importance of good initialization to active contours

In this subsection, the necessity for a good initialization in the proposed affine active contour for segmentation is elucidated.

In Figure 6.13, it is shown that the object segmentation result by using the affine active contour is bad given an initial contour far from the object, whereas the object segmentation result is good given the initial contour shape well-aligned with the object in the image. All these elucidate that a good initialization is

Figure 6.11: The significance of the affine invariance of the interior-points-to-shape relation. The top row shows the shapes generated by the affine-invariant model in Eq. (6.8). The bottom row shows the shapes generated by the model without affine invariance in Eq. (6.7). The blue dots are the interior feature points, the red curves are the implicit contours, and the black circles are the explicitly transformed contours.

crucial to the conventional active contours which are based on gradient descent curve evolution. Note that the results shown in Figure 6.13 are slightly different from those shown in Figure 6.1. This is because the two methods converge to different local optimal solutions when given slightly different initial conditions.

6.4.3 The fundamental problem of global search

The global optimization for active contours have been considered in [89, 90]. The main idea behind these works is to construct efficient approximations to the exhaustive search. This subsection presents the results of exhaustive search of the

Figure 6.12: The normalized Hausdorff distances between the shapes implicitly defined by the transformed interior points and the original shapes (horizontal) v.s. the distance between the explicitly transformed contour shapes and the original shape (vertical).

shape of interest in a query image. Exhaustive search is performed for both the Chan-Vese model and the GAC model. In the implementation, the energy of the Chan-Vese model or the GAC model is evaluated for 8 orientations of the given shape, i.e. $\{0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}, \pi, \frac{5\pi}{4}, \frac{3\pi}{2}, \frac{7\pi}{4}\}$, at all possible locations within the image domain. Then, the location and orientation corresponding to the smallest energy is chosen as the global optimal solution. When the size of the shape is known and the image is simple, the search can provide acceptable results, as shown in Figure 6.14. Note that, knowing the size is crucial in active contours, since active contours all prefer small shapes. Such preference is understood as the Minimum Description Length principle. Without specifying the size in advance, the exhaustive search may provide degenerate solutions.

More often than not, the global search does not provide a reliable solution to the object segmentation problem in general real images even when the size is known. This is because the object of interest does not necessarily correspond to the global optimal solution of the formulated energy minimization problem. Such a result is shown in Figure 6.15. On the one hand, it can be observed that the global optimal solution to the Chan-Vese model, the contour at the top-left corner, provides the region in which the image values contrast most significantly with the outer region.

Figure 6.13: The object segmentation by the affine active contour driven by affine motion of interior points. (a) is a bad initialization which leads to a bad result shown in (b). (c) displays a good initialization which leads to a good result shown in (d).

On the other hand, the GAC model locates the group of edges that fit the shape best but away from the graffiti face. Neither of the results are satisfactory.

Figure 6.14: The object segmentation by exhaustive search in simple image. (a) is the result by Chan-Vese model (b) is the result by GAC model

Figure 6.15: The object segmentation by exhaustive search in a cluttered image. (a) shows the result by the Chan-Vese model (b) shows the result by the GAC model

6.4.4 Automatic object segmentation by matching-constrained active contour

The experiments presented in this subsection evaluate the proposed matching-constrained active contour. The centroids of the template objects in the matching are taken as the origin $(0,0)$. This position can be anywhere, and it does not affect the feature point matching. This is because the formulation of feature point matching does not involve the absolute position.

From the point correspondences (obtained from the SIFT matching) shown in the left-most columns of Figures 6.16, 6.17 and 6.18. The initial contours are generated and they shown in the middle columns of the figures, and the final segmentation results are shown in the right-most columns of the figures. It can be observed that the initial contours produced by the proposed interior-points-to-shape relation are close to the object boundaries, and the active contour further improves the boundary location. The proposed method is capable of achieving the desirable segmentation results. The initial and final segmentation results are compared with manually labeled ground truth shapes by Jaccard similarity measure to validate the reported visual observations quantitatively. The results are summarized in Table 6.2. The region enclosed by the final contour overlaps well

with the ground truth. The computational costs of the entire object segmentation framework on the images are also presented in Table 6.3 for reference (Note that the entire process is run in Matlab on a Core-i5 2.4Ghz PC). The large computational cost was due to the difficulty in determining the convergence by judging on the changes of contours between successive iterations. The evolving contour was in fact oscillating. The contour evolution was sometimes terminated when a predetermined maximum number of iterations was reached.

Table 6.2: The Jaccard similarity against the ground truth.

	Figure 6.16				Figure 6.17		Figure 6.18
	(a)	(b)	(c)	(d)	(a)	(b)	
Initial	0.73	0.75	0.74	0.63	0.53	0.68	0.65
Final	0.75	0.84	0.86	0.86	0.90	0.93	0.75

Table 6.3: Computational costs (in seconds)

Figures	6.16(a)	6.16(b)	6.16(c)	6.16(d)	6.17(a)	6.17(b)	6.18
Size(pixel)	320×400	320×400	320×400	320×400	296×448	282×448	411×408
M&R	17.83	18.3	18.53	17.9	18.44	16.27	10.05
GC	4.26	4.21	4.20	4.32	7.14	7.10	2.29
AC	98.14	40.93	82.03	650.56	504.97	401.95	82.95
Total	120.23	63.46	104.75	672.78	530.56	425.32	95.29

M&R=matching and registration, GC= generation of the initial contour, AC = Active contour before convergence, Total = Total running time

It has been implicitly assumed that a better shape model leads to a better result of segmentation throughout the chapter. This can be ascertained by experiments. In the experiment, the 50 shape models due to the 50 random σ values generated previously were used for extraction of the leaf in the top image in Figure 6.17. The results are measured quantitatively by the Jaccard similarity. Figure 6.19 indicates the strong correlation between the quality of shape modeling, in terms of fitting score, and the quality of segmentation, in terms of the Jaccard similarity between the result and the ground truth.

Figure 6.16: Segmentation of an object of different poses in noisy images in the presence of occlusion by leaves, clutter background and unclear boundary on the top. The rows of the figures from top to bottom are referred to as (a) (b) (c) and (d).

6.4.5 Discussions

There is a tolerance level τ in the formulation of the matching-constrained active contour to be determined by the user. Here, some pointers are provided to help users select their appropriate τ . Since a small τ may only allow small deviation

Figure 6.17: Segmentation of an object of relatively complex shape under different poses and scales with non-ideal initial matchings. The top and bottom triple of figures are labeled as (a) and (b).

Figure 6.18: Segmentation of an object of complex shape under different scale, clutter background and indefinite boundary.

from the initial solution even when the boundary is not very close to the initial contour, one may not need to constrain the optimization intensively. In the images for evaluation, the boundaries of the objects of interest correspond to the local optimal solutions to the active contour energy functional. Hence, only the gradient descent given the initial feasible solution produced by optimal matching is needed without compromising the solution to the constraint by projected gradient descent. The constraint should be more useful during the search process for the optimal

Figure 6.19: Quality of segmentation (vertical) v.s. score of model fitting (horizontal)

solution when the boundary is actually not a local optimal solution to the active contour model.

Objects may deform both rigidly and non-rigidly. In this chapter, a general type of deformation, i.e. the affine transformation, is also considered. The more general non-rigid deformations can be estimated by following the nonrigid shape prior modeling, e.g. [35, 72, 86], once the affine shape transformation is known.

Robust detection of the convergence of the gradient descent based method is a research problem in many areas such as optimization and machine learning. In the experiments, the algorithm is terminated when the average change of the affine parameters is less than a threshold or when the maximum iteration is reached.

The computational cost depends on the number of selected feature points on the template object. A careful selection of the feature points can reduce the computational cost drastically while preserving the segmentation accuracy.

6.5 Concluding remarks

In this chapter, a new constrained optimization base formulation of active contours model, with the initial contour provided by object matching, is proposed. The contributions of this work lie in several aspects. First, a mathematical model of

the affine invariant relation between the interior feature points and the continuous contour shape is determined, which enables the shape alignment based on point alignment. Second, a novel affine active contour model driven by the affine motion of the interior points is proposed. Third, the points-to-shape alignment and affine active contour are unified through a constrained optimization model, giving rise to a mathematical model of automatic object segmentation. It is demonstrated that the method achieves automatic object segmentation under various real scenes. This proposed work expands the capacity of the active contour approach for object segmentation, and the extension has practical significance in that the initial curve no longer needs to be provided by the user.

Chapter 7

Conclusion and Recommendations

7.1 Conclusion

This thesis addresses three fundamental problems faced in active contour based object segmentation methods.

- (I) How can we formulate active contour models for segmenting objects based on its boundary properties or its regional properties?
- (II) If the object boundary or the object region appears partially in the image, due to occlusion or cluttered scene, how can we reformulate the active contour models in order to segment the object and recover the original object shape?
- (III) Given a real image containing a clutter of objects, how can we incorporate the prior information of the object of interest in the active contour formulation for segmenting the object of interest correctly without user intervention?

To address Problem (I), two contributions were made as presented in Chapters 3 and 4. To address Problem (II), a contribution is made as presented in Chapter 5. To address Problem (III), a contribution was made as presented in Chapter 6.

These contributions are summarized hereafter.

Chapter 3: GeoSnakes and equilibrium flow - A well-known problem with the edge-based active contours, namely the problem of early termination of the curve evolution, is addressed. The problem is studied by examining the fundamental formulation of edge-based active contour models in which only the normal component of the gradient of the edge indicator is considered. It is argued that the tangential component should also be included. To include the tangential component, a new edge-based active contour model called the GeoSnakes is formulated, and the numerical solution to the GeoSnakes is based on a new type of curve evolution named the equilibrium flow.

Chapter 4: Piecewise linear patch reconstruction - For segmentation of non-smooth structures, e.g. textures, the respective image features have previously been used in the active contour methods as a separate process. In this thesis, the segmentation of non-smooth image structures is formulated as a novel active contour model in terms of piecewise linear patch reconstruction, motivated by the Mumford-Shah model. To avoid heuristics, it is further proven that the minimization of the patch reconstruction error leads to reduction of the error of segmentation. Without a separate process of feature extraction, the segmentation is accomplished in a principled way by minimizing the proposed single energy functional of piecewise linear patch reconstruction through curve evolution and optimal linear patch reconstruction.

Chapter 5: Prior variation shape evolution - Due to occlusion and cluttered background resulting in missing parts or overlapping objects in the image, the energy minimization of the generic active contour models by curve evolution cannot extract the object shape of interest correctly. An object of interest may

also undergo some deformation causing changes to its observed shape. Depending on the type of the object the deformation may be rigid or non-rigid. The third contribution of this thesis is a novel deterministic model of rigid and nonrigid shape deformations based on curve evolution, giving rise to the Prior Variation Shape Evolution (PVSE) equation. The PVSE equation is used to minimize active contour energy for locating the object boundary of the deformed and/or recovered shape.

Chapter 6: Matching-constrained active contour - To segment the object of interest, an active contour is often initialized manually by an user. By incorporating local feature matching and shape prior modeling into the active contour framework, a matching-constrained active contour model for automatic object segmentation is proposed. The initial contour, which is the initial feasible solution to the constrained optimization problem, is generated by local feature matching followed by an affine point-to-shape alignment.

7.2 Recommendations for future research

Regarding the Problem (I) addressed in Chapters 3 and 4, there are still fundamental questions unanswered. The original edge-based and region-based active contour models are based on different principles of segmentation. It is well-known that the two kinds of models are suitable for different images. Therefore, the unification of the two models would be beneficial. Recently, there exist research works on unifying the edge-based and region-based active contour models, such as [39, 40]. However, an outstanding question is that how we can judge the unification of the two models. This gives rise to a series of questions. For example, we may need to know what is the hidden relationship between the two models that may possibly be used for the unification. The answer to this question may lie in the fundamental

research of cognitive vision.

In Problem (II) addressed in Chapter 5, it is required to get the shape of the object even when the shape of the object does not appear properly in the image. This problem is highly ill-posed since the shape cannot be recovered without knowing the shape information *a priori*. Therefore, an important question is how to model the prior shape which agrees with the shape in the image. This is the question that many research papers [31–33, 35] and the work presented in Chapter 5 try to answer. However, the answers to this question appear to be far from a consensus. For example, regarding the work presented in Chapter 5, though the deterministic model may be ready for use without training, the statistical model can be more reliable for modeling the shape variations of a given object. The important question is how to balance the requirement of large sample and the model error.

The Problem (III) addressed in Chapter 6 is also worthy of further elaboration. In the present approach, the feature point matching with a single template is used to generate the initial contour. There are still problems with the feature point matching approach, e.g. partial matching, variation of appearance, and quantitative evaluation. Besides, the present mathematical model of boundary estimation by matching the interior feature points is based on the affine invariance. However, the interior points may not undergo affine transformation. Hence, more advanced mathematical model of the boundary estimation should be considered. Furthermore, the accuracy of the model should be quantified in modeling and evaluation.

Appendices

A Gradient descent method for active contours

A.1 Functional gradient and gradient descent

The solution to an active contour model is an infinite dimensional curve. The discretization will cause indefinite number of dimensions of the solution. This problem is *hard* in its nature to be solved via exhaustive search. For minimizing the active contour energies in the form of contour integrals, the *gradient descent* method is often used to update the evolving contour curve until convergence.

Active contour model is a special case of the functional energy minimization problem, i.e. that of the minimization of integrals by finding an optimal function. If functions can be considered as points in the abstract spaces of functions, such as Banach spaces, Hilbert spaces, or the Reproducing Kernel Hilbert Spaces, etc., the mappings from the space of functions to values are called the functionals. Bounded scalar-valued functionals can be maximized or minimized with respect to the input function by generalizing the mathematical setting of finite dimensional optimization. To maximize or minimize the functional mappings, we may mimic the case in a finite-dimensional space. Informally, if we consider to minimize the energy functional $E : \mathcal{F} \mapsto \mathbb{R}^+$, where $\mathcal{F}$ is a space of differentiable functions. The minimal order of differentiability is given by the highest order of differentiation involved in the functional energy. The necessary condition of a function $f^* \in \mathcal{F}$ to

be the optimal solution can be established as follows:

$$\left. \frac{dE}{d\epsilon} \right|_{f=f^*} = \lim_{\epsilon \rightarrow 0} \frac{E(f^* + \epsilon\eta) - E(f^*)}{\epsilon} = 0, \text{ for all } \eta. \quad (\text{A.1})$$

According to the theory of *Fréchet* derivative, there exists a bounded linear operator ∇E satisfying the following:

$$\left. \frac{dE}{d\epsilon} \right|_{f=f^*} = \lim_{\epsilon \rightarrow 0} \frac{E(f^* + \epsilon\eta) - E(f^*)}{\epsilon} = \nabla E(f^*) \cdot \eta, \quad (\text{A.2})$$

where “ $\cdot$ ” denotes the dot product of functions. The above linear approximation of the limit was first proposed by *Gâteaux*. Hence, it is called the *Gâteaux* derivative. The operator ∇E is actually the functional gradient. This identity leads to the fact that $E(f^*) = 0$. Besides, the functional gradient ∇E can be used for the gradient descent method. Let us consider an artificial parametrization of the function f with respect to time, we have the following property for minimizing the energy:

$$\begin{aligned} \frac{d(E(t))}{dt} &= \lim_{\Delta t \rightarrow 0} \frac{E(f(t + \Delta t)) - E(f(t))}{\Delta t} \\ &= \lim_{\Delta t \rightarrow 0} \frac{E(f(t) + \Delta t \frac{\partial f}{\partial t}) - E(f(t))}{\Delta t} \\ &= \nabla E \cdot \frac{\partial f}{\partial t}. \end{aligned} \quad (\text{A.3})$$

We may note that if we choose

$$\frac{\partial f}{\partial t} = -\nabla E,$$

we can decrease the functional energy of $E(f)$.

In this section, it is assumed that the gradient of the energy with respect to the curve $C(p)$ has been obtained. By using gradient descent, one can obtain the

Figure A.1: Contour splitting.

gradient descent equation that governs the curve evolution in a form as follows:

$$\frac{\partial C(p)}{\partial t} = -\nabla E(C(p)) = \begin{cases} \frac{\partial x(p)}{\partial t} = -E_x(C(p)) \\ \frac{\partial y(p)}{\partial t} = -E_y(C(p)) \end{cases}. \quad (\text{A.4})$$

This equation is the curve evolution equation for energy minimization.

In the original framework of active contours, splines were used to represent the curve. Hence, the infinite dimensional partial differential equations of curve evolution are solved by updating the knot points of the splines. However, there is a major drawback of this formulation. The spline curve is continuous. If the initial curve is closed, then the convergent curve will still be one closed curve. Consequently, this method cannot be used to segment complex images with unknown number of objects. Specifically, the curve evolution of contour splitting (or merging) as shown in Figure A.1 cannot happen.

This technical problem in the area of computational physics was tackled by using an implicit representation of the closed curves [14]. The idea is to use the zero level set of a two dimensional function ϕ to represent a two dimensional curve C , which is called the level set method. This is illustrated in Figure A.2. One can observe that the topology of the curve changes, i.e., a closed curve is changed to be two closed curves or vice versa. This technique was “transplanted” to the areas of image processing and computer vision as a result of a series of research efforts such as [15–17, 24, 96, 97]. This has strong influence on subsequent work in active

Figure A.2: The idea of the level set method. The dotted curves on ϕ are the zero level sets

contour.

It is now possible to detail the formulation of this elegant idea of level set method for realizing curve evolution [13]. Since the curve $C(p, t)$ can be represented by the zero level set of the signed distance function $\phi(p, t)$ defined in Eq. (2.7), we have the following:

$$C(p, t) = \{x, y | \phi(x, y, t) = 0\}. \quad (\text{A.5})$$

where the parameter t denotes the evolving time.

Taking derivative of Eq. (A.5), we obtain the following:

$$\begin{aligned} \left. \frac{d\phi}{dt} \right|_{\phi=0} &= \left. \frac{\partial \phi}{\partial t} + \nabla \phi \cdot \vec{V} \right|_{\phi=0} = 0 \\ \Rightarrow \left. \frac{\partial \phi}{\partial t} \right|_{\phi=0} &= -\nabla \phi \cdot \vec{V} \Big|_{\phi=0}, \end{aligned} \quad (\text{A.6})$$

where

$$\begin{aligned}
 \vec{V}\Big|_{\phi=0} &= \left[\frac{dx}{dt}, \frac{dy}{dt} \right]^T \Big|_{\phi=0} = \frac{dC}{dt} \\
 &= \frac{\partial C}{\partial t} + \frac{\partial C}{\partial p} \frac{dp}{dt} \\
 &= \frac{\partial C}{\partial t} + \frac{dp}{dt} \left\| \frac{\partial C}{\partial p} \right\| \vec{T},
 \end{aligned} \tag{A.7}$$

and $\vec{T}$ is the tangent of the curve. Obviously, $\vec{T} \cdot \vec{N} = 0$ and $\frac{\nabla \phi}{\|\nabla \phi\|} = -\vec{N}$, and $\vec{N}$ is the outward contour normal.

Hence, combining with the general formulation of curve evolution Eq. (A.4) we obtain

$$\begin{aligned}
 \frac{\partial \phi}{\partial t} \Big|_{\phi=0} &= -\nabla \phi \cdot \vec{V} \Big|_{\phi=0} \\
 &= -\nabla \phi \cdot \left(\frac{\partial C}{\partial t} + \frac{dp}{dt} \left\| \frac{\partial C}{\partial p} \right\| \vec{T} \right) \\
 &= -\nabla \phi \cdot \frac{\partial C}{\partial t},
 \end{aligned} \tag{A.8}$$

which is a variant of the level set equation for curve evolution. Note that the curve evolution governed by the dynamic of $\frac{\partial C}{\partial t}$ have the following general form:

$$\frac{\partial C}{\partial t} = \alpha \vec{N} + \beta \vec{T}, \tag{A.9}$$

where $\alpha = -\langle \nabla E(C), \vec{N} \rangle$ and $\beta = -\langle \nabla E(C), \vec{T} \rangle$. The level set equation in Eq. (A.8) can be simplified if ϕ is the signed distance function defined in Eq. (2.7), where $\|\nabla \phi\| = 1$ almost everywhere.

$$\begin{aligned}
 \frac{\partial \phi}{\partial t} \Big|_{\phi=0} &= -\nabla \phi \cdot \frac{\partial C}{\partial t} \\
 &= \alpha \|\nabla \phi\|.
 \end{aligned} \tag{A.10}$$

By solving the level set equation with an initial ϕ_0 until convergence at ϕ^* , one obtains the convergent solution to the gradient descent optimization, represented by $\phi^* = 0$. The level set equation Eq. (A.10), which is in the category of Hamilton-Jacobian Equations in the literature of partial differential equations, is then solved numerically. Efficient numerical implementation of this kind of equations has also been developed [13, 53]. The implementation basically involves two steps. One is the finite difference version of the level set equation in Eq. (A.10). The other is the re-initialization to maintain the function ϕ to be the signed distance function. The details on this is omitted in this thesis. Interested readers are referred to [12, 13, 53].

Interestingly, due to the more advanced reformulation of the Mumford-Shah model in (2.6) initiated by Chan and Vese in [24], most of the existing region-based active contours can be minimized by the level set method without the need of converting the region integrals to contour integrals followed by derivation of the energy-minimizing curve evolution equation and application of the level set method.

A.2 The calculus of variations for active contours

Needless to say, the actual derivation of the linear operator ∇E is of great importance for the optimization. The derivation is now known as the calculus of first variations, or the *calculus of variations* if there is no risk of confusion. The η in Eq. (A.1) is the variation of the function f^* .

One may refer to Chapter 21 of Peter J. Olver's notes of Applied Mathematics [98] for a fundamental introduction of this topic. When going through the notes, the first two chapters of [99] may also be of interest. This technique of calculus of variations has been widely used in the area of optimal shape design of mechanical parts. Some fundamentals of optimal shape design can be found in [100].

A simplest case of the calculus of variations can be shown as follows:

$$f^* = \underset{f}{\operatorname{argmin}} J(f) = \int_{\Omega} F(f(x), f_x(x)) dx. \quad (\text{A.11})$$

The subscript denotes the first derivative, F is a function of $f(x), f_x(x)$, Ω is the integral domain of x , and $f_x(x)$ is the first order derivative of f w.r.t. x . This formulation seems to be rather abstract; however, it can be shown that the gradient of $J(f)$ w.r.t. f can be more or less easily derived by following a calculus of directional derivatives similar to the derivation in finite dimension. In other words, the idea is to obtain the gradient by using directional derivative on an arbitrary direction η . By assuming $\Omega = [0, 1)$, and given the boundary condition $\eta|_{\{0,1\}} = 0$, one can have the following:

$$\begin{aligned} \left. \frac{dJ(f^* + \epsilon\eta)}{d\epsilon} \right|_{\epsilon=0} &= \int_{\Omega} \frac{\partial F(f^*, f_x^*)}{\partial f} \eta + \frac{\partial F(f^*, f_x^*)}{\partial f_x} \eta_x dx \\ &= \int_{\Omega} \frac{\partial F(f^*, f_x^*)}{\partial f} \eta dx + \underbrace{\frac{\partial F(f^*, f_x^*)}{\partial f_x} \eta \Big|_{\partial\Omega=\{0,1\}}}_{=0} - \int_{\Omega} \frac{d}{dx} \frac{\partial F(f^*, f_x^*)}{\partial f_x} \eta dx \\ &\quad \text{(by integration by parts)} \\ &= \int_{\Omega} \left\{ \frac{\partial F(f^*, f_x^*)}{\partial f} - \frac{d}{dx} \frac{\partial F(f^*, f_x^*)}{\partial f_x} \right\} \eta dx \\ &= \int_{\Omega} \nabla J(f) \eta dx \\ &= \left\langle \nabla J(f), \eta \right\rangle_{L_2}, \end{aligned} \quad (\text{A.12})$$

where η is the so-called *variation* on f^* , $\langle \cdot, \cdot \rangle_{L_2}$ denotes the L_2 inner product of univariate functions,

$$\nabla J(f) = \frac{\partial F(f^*, f_x^*)}{\partial f} - \frac{d}{dx} \frac{\partial F(f^*, f_x^*)}{\partial f_x}$$

is the gradient of the functional $J(f)$. In the derivation above, we have assumed that the boundary value of the variation η is zero. We will assume the boundary values of the variation to be zero in the rest of the thesis.

Further, by letting the negative gradient be zero, we obtain the Euler-Lagrange equation. This technique of calculus of the first order derivative of functional energies is named the Calculus of Variation which is the actual derivation of the *Gâteaux* derivative for functionals in the integral form. This derivation can be easily generalized to more complicated cases, such as the following:

$$f^* = \underset{f}{\operatorname{argmin}} J(f) = \int_{\Omega} F(f(\vec{x}), f_{x_1}(\vec{x}), f_{x_2}(\vec{x}), \dots, f_{x_i x_j, \dots}(\vec{x})) dx_1 dx_2 \dots$$

By applying the multivariate version of integration by parts, i.e., the Green's identity, the *Gâteaux* derivative will be of the form below:

$$\begin{aligned} \left. \frac{dJ(f^* + \epsilon\eta)}{d\epsilon} \right|_{\epsilon=0} &= \int_{\Omega} \nabla J(f) \eta dx_1 dx_2 \dots \\ &= \left\langle \nabla J(f), \eta \right\rangle_{L_2}, \end{aligned} \quad (\text{A.13})$$

where $\left\langle \cdot, \cdot \right\rangle_{L_2}$ denotes the L_2 inner product of multivariate functions.

It is now possible to show how this technique can be used in the framework of active contours.

Example: Snakes [3] The *Gâteaux* derivative of the snake functional energy in Eq. (2.1) can be written as follows:

$$\begin{aligned} \left. \frac{dE(C + \epsilon\eta)}{d\epsilon} \right|_{\epsilon=0} &= \int [\alpha C'(p)^T \eta' + \beta C''(p)^T \eta''] - \nabla E_{ext}(C(p))^T \eta dp \\ &= \int [-\alpha C_{pp}(p) + \beta C_{ppp}(p) - \nabla E_{ext}(C(p))]^T \eta dp \\ &= \left\langle \nabla E, \eta \right\rangle_{L_2^*}, \end{aligned} \quad (\text{A.14})$$

where $\langle \cdot, \cdot \rangle_{L_2^*}$ denotes the L_2 inner product of vector functions, and C_{pp} and C_{pppp} are the second and fourth order partial derivatives of $C(p, t)$ w.r.t. p . As stated in Eq. (A.4), the gradient descent w.r.t. the curve C can be written as follows:

$$\frac{\partial C}{\partial t} = \alpha C_{pp}(p) - \beta C_{pppp}(p) + \nabla E_{ext}(C(p)). \quad (\text{A.15})$$

★ End of example ★

Example: Geodesic Active Contour (GAC) [6, 17] The *Gâteaux* derivative of the GAC energy in (2.2) can be written as follows:

$$\begin{aligned} & \left. \frac{d}{d\epsilon} \int_0^L g(C(p) + \epsilon \eta(p)) \|C_p + \epsilon \eta_p\| dp \right|_{\epsilon=0} \\ &= \int_0^L \langle \nabla g \|C_p\|, \eta \rangle + g \left\langle \frac{C_p}{\|C_p\|}, \eta_p \right\rangle dp \\ &= \int_0^L \langle \nabla g \|C_p\|, \eta \rangle - \langle \nabla g, C_p \rangle \left\langle \frac{C_p}{\|C_p\|}, \eta \right\rangle - g \left\langle \frac{d}{dp} \left(\frac{C_p}{\|C_p\|} \right), \eta \right\rangle dp \\ &= \int_0^L \langle \nabla g \|C_p\|, \eta \rangle - \langle \nabla g, \frac{C_p}{\|C_p\|} \rangle \left\langle \frac{C_p}{\|C_p\|} \right\rangle \|C_p\| \eta - g \left\langle \kappa \vec{N} \frac{ds}{dp}, \eta \right\rangle dp \\ &= \int_0^L \langle \nabla g \vec{N}, \eta \rangle \|C_p\| - g \left\langle \kappa \vec{N}, \eta \right\rangle \|C_p\| dp \\ &= \int_0^{L(C)} \langle \nabla g \vec{N} - g \kappa \vec{N}, \eta \rangle ds, \end{aligned} \quad (\text{A.16})$$

where we applied a series of identities: $\vec{T} = \frac{C_p}{\|C_p\|}$, $\frac{d\vec{T}}{dp} = \kappa \vec{N}$ and $ds = \|C_p\| dp$ as well as the integration by parts, κ is the curvature of the curve.

Thus, the curve evolution equation for GAC is as follows:

$$\frac{\partial C}{\partial t} = -\nabla E = g \kappa \vec{N} - \langle \nabla g, \vec{N} \rangle \vec{N}. \quad (\text{A.17})$$

★ End of example ★

In the area of optimal design [100], it has been identified that the calculus of variations cannot be easily applied to the minimization of the type of the integrals

in Eqs. (2.4) and (2.5) shown previously. These integrals are called the domain integrals in optimal design. The objective is to find the optimal domain on which the integral tends to be maximal or minimal. However, the direct calculus of variations is not applicable to this case, since it is difficult to represent the changes in the integral due to the variations of the domain. Nevertheless, Zhu and Yuille [18] established the equivalence between the domain integral of Eq. (2.5) and a form of contour integral.

Example: Region Competition [18] As in [18], the functional for the region integral in the region competition can be written in a general form as follows:

$$E(\{c_l\}, \{\Omega_l\}) = \int_{\Omega_l} \mathcal{E}_l dx dy + \int_{\mathcal{D}/\Omega_l} \mathcal{E}_{\bar{l}} dx dy. \quad (\text{A.18})$$

The following identity connects the two types of integrals:

$$\iint_{\Omega} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy = \int_{\partial\Omega} P dx + Q dy = \int_0^L P x_s ds + Q y_s ds, \quad (\text{A.19})$$

which is the Green's theorem as a special case of the Stokes' theorem.

Let

$$Q(x, y) = \frac{1}{2} \int_0^x E_l(t, y) dt, \quad P(x, y) = -\frac{1}{2} \int_0^y E_l(x, t) dt,$$

we have the following:

$$\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = E_l(x, y).$$

Substituting the above into (A.19), we obtain the contour integral form of the

region-based active contours as follows:

$$\begin{aligned}
& \iint_{\Omega_l} E_l dx dy \\
&= \frac{1}{2} \int_0^L \left[\int_0^{y(s)} E_l(x(s), t) dt, - \int_0^{x(s)} E_l(t, y(s)) dt \right] [x_s, y_s]^T ds \\
&= \frac{1}{2} \int_0^L \left[\int_0^{y(s)} E_l(x(s), t) dt, - \int_0^{x(s)} E_l(t, y(s)) dt \right] \cdot \vec{T} ds,
\end{aligned} \tag{A.20}$$

where $\vec{T} = [x_s, y_s]^T$, and the default direction is anti-clockwise.

The *Gâteaux* derivative of the contour integral is the following:

$$\begin{aligned}
& \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} \left\{ \frac{1}{2} \int_0^L \left(\int_0^{y(s)+\epsilon v(s)} E_l(x(s) + \epsilon u(s), t) dt \right) (x_s + \epsilon u_s) ds \right. \\
& \quad \left. - \frac{1}{2} \int_0^L \left(\int_0^{x(s)+\epsilon u(s)} E_l(t, y(s) + \epsilon v(s)) dt \right) (y_s + \epsilon v_s) ds \right\} \\
&= \frac{1}{2} \int_0^L \left\{ \left(\int_0^{y(s)} \frac{\partial E_l(x(s), t)}{\partial x} dt \right) u(s) + E_l(x(s), y(s)) v(s) \right\} x_s ds \\
& \quad + \frac{1}{2} \int_0^L \left(\int_0^{y(s)} E_l(x(s), t) dt \right) u_s ds \\
& \quad - \frac{1}{2} \int_0^L \left\{ \left(\int_0^{x(s)} \frac{\partial E_l(t, y(s))}{\partial y} dt \right) v(s) + E_l(x(s), y(s)) u(s) \right\} y_s ds \\
& \quad - \frac{1}{2} \int_0^L \left(\int_0^{x(s)} E_l(t, y(s)) dt \right) v_s ds.
\end{aligned} \tag{A.21}$$

By integration by parts,

$$\begin{aligned}
& \int_0^L \left(\int_0^{y(s)} E_l(x(s), t) dt \right) u_s ds \\
&= - \int_0^L \frac{\partial}{\partial s} \left(\int_0^{y(s)} E_l(x(s), t) dt \right) u ds \\
&= - \int_0^L \left(\int_0^{y(s)} \frac{\partial E_l(x(s), t)}{\partial x} dt \right) x_s u ds - \int_0^L E_l(x(s), y(s)) y_s u ds,
\end{aligned} \tag{A.22}$$

$$\begin{aligned}
& \int_0^L \left(\int_0^{x(s)} E_l(t, y(s)) dt \right) v_s ds \\
&= - \int_0^L \left(\int_0^{x(s)} \frac{\partial E_l(t, y(s))}{\partial y} dt \right) y_s v ds - \int_0^L E_l(x(s), y(s)) x_s v ds.
\end{aligned} \tag{A.23}$$

To sum up, Eq. (A.21) becomes the following:

$$\begin{aligned}
& \frac{d}{d\epsilon} \Big|_{\epsilon=0} \left\{ \frac{1}{2} \int_0^L \left(\int_0^{y(s)+\epsilon v(s)} E_l(x(s) + \epsilon u(s), t) dt \right) (x_s + \epsilon u_s) ds \right. \\
& \quad \left. - \frac{1}{2} \int_0^L \left(\int_0^{x(s)+\epsilon u(s)} E_l(t, y(s) + \epsilon v(s)) dt \right) (y_s + \epsilon v_s) ds \right\} \\
&= \int_0^L E_l(x(s), y(s)) x_s v ds - \int_0^L E_l(x(s), y(s)) y_s u ds \\
&= \int_0^L E_l(x(s), y(s)) \langle \vec{N}, [u, v]^T \rangle ds,
\end{aligned} \tag{A.24}$$

where $\vec{N}$ is the outward normal.

Likewise,

$$\begin{aligned}
& \iint_{\mathcal{D}/\Omega_l} E_{\bar{l}} dx dy \\
&= \frac{1}{2} \int_L^0 \left[\int_0^{y(s)} E_{\bar{l}}(x(s), t) dt, - \int_0^{x(s)} E_{\bar{l}}(t, y(s)) dt \right] [x_s, y_s]^T ds \\
&= \frac{1}{2} \int_L^0 \left[\int_0^{y(s)} E_{\bar{l}}(x(s), t) dt, - \int_0^{x(s)} E_{\bar{l}}(t, y(s)) dt \right] \cdot \vec{T} ds,
\end{aligned} \tag{A.25}$$

and,

$$\begin{aligned}
& \frac{d}{d\epsilon} \Big|_{\epsilon=0} \left\{ \frac{1}{2} \int_L^0 \left(\int_0^{y(s)+\epsilon v(s)} E_{\bar{l}}(x(s) + \epsilon u(s), t) dt \right) (x_s + \epsilon u_s) ds \right. \\
& \quad \left. - \frac{1}{2} \int_L^0 \left(\int_0^{x(s)+\epsilon u(s)} E_{\bar{l}}(t, y(s) + \epsilon v(s)) dt \right) (y_s + \epsilon v_s) ds \right\} \\
&= \int_L^0 E_{\bar{l}}(x(s), y(s)) x_s v ds - \int_L^0 E_{\bar{l}}(x(s), y(s)) y_s u ds \\
&= \int_L^0 E_{\bar{l}}(x(s), y(s)) \langle \vec{N}, [u, v]^T \rangle ds = \int_0^L E_{\bar{l}}(x(s), y(s)) \langle -\vec{N}, [u, v]^T \rangle ds.
\end{aligned} \tag{A.26}$$

Thus, the curve evolution equation corresponding to the region integral in (A.18) is the following:

$$\frac{\partial C}{\partial t} = \left(E_{\bar{l}}(x(s), y(s)) - E_l(x(s), y(s)) \right) \vec{N}. \quad (\text{A.27})$$

The full model of region competition often contains a term penalizing the contour complexity in terms of contour length. The full model with the term of contour complexity is as follows:

$$E(\{c_l\}, \{\Omega_l\}) = \int_{\Omega_l} \mathcal{E}_l dx dy + \int_{\mathcal{D}/\Omega_l} \mathcal{E}_{\bar{l}} dx dy + \nu \int_{\partial\Omega_l} \|C_p\| dp. \quad (\text{A.28})$$

To minimize this term together with the region integral, we may derive the functional gradient of this term separately. The derivation of the functional gradient may follow the previous derivations for GAC for the weight of $g = 1$. Therefore, the full curve evolution equation for region competition is the following:

$$\frac{\partial C}{\partial t} = \left(E_{\bar{l}}(x(s), y(s)) - E_l(x(s), y(s)) \right) \vec{N} + \nu \kappa \vec{N}. \quad (\text{A.29})$$

★ End of example ★

The previous three active contour models can be implemented via the direct finite difference discretization of the derived curve evolution equations. The curve evolution can be implemented via the more advantageous level set method. Due to the reformulation by Chan and Vese [24], the energy minimization of the region competition model leads to the direct level-set form of curve evolution, and the derivation of the *Gâteaux* derivative is much simpler than that of the original region competition. In the following, the idea of [24] is presented.

Example: Chan-Vese model [24] The reformulation of the region competition

by Chan and Vese is the following:

$$E(\{c_l\}, \{H_l\}) = \int_{\mathcal{D}} \mathcal{E}_l H_l dx dy + \int_{\mathcal{D}} \mathcal{E}_{\bar{l}} (1 - H_l) dx dy + \nu \int_{\partial\Omega} \delta(\phi) \|\nabla\phi\| dx dy, \quad (\text{A.30})$$

where $H_l = H(\phi_l)$, H is a Heaviside function, ϕ_l is the signed distance function. A similar form of the above has been presented in (2.6). It can be proven that (A.30) is identical to (A.28).

The *Gâteaux* derivative of (A.30) is as follows:

$$\begin{aligned} & \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} E(\phi_l + \epsilon\psi) \\ &= \int_{\mathcal{D}} \mathcal{E}_l \delta(\phi) \psi dx dy - \int_{\mathcal{D}} \mathcal{E}_{\bar{l}} \delta(\phi) \psi dx dy \\ &+ \nu \int_{\partial\Omega} \frac{d\delta(\phi)}{d\phi} \psi \|\nabla\phi\| + \delta(\phi) \frac{\nabla\phi \cdot \nabla\psi}{\|\nabla\phi\|} dx dy \\ &= \int_{\mathcal{D}} \mathcal{E}_l \delta(\phi) \psi dx dy - \int_{\mathcal{D}} \mathcal{E}_{\bar{l}} \delta(\phi) \psi dx dy \\ &+ \nu \int_{\partial\Omega} \left\langle \frac{d\delta(\phi)}{d\phi} \nabla\phi, \psi \nabla\phi \right\rangle + \delta(\phi) \frac{\nabla\phi \cdot \nabla\psi}{\|\nabla\phi\|} dx dy \\ &= \int_{\mathcal{D}} \mathcal{E}_l \delta(\phi) \psi dx dy - \int_{\mathcal{D}} \mathcal{E}_{\bar{l}} \delta(\phi) \psi dx dy \\ &- \nu \int_{\partial\Omega} \nabla \cdot (\psi \nabla\phi) \delta(\phi) + \delta(\phi) \frac{\nabla\phi \cdot \nabla\psi}{\|\nabla\phi\|} dx dy \\ &= \int_{\mathcal{D}} \mathcal{E}_l \delta(\phi) \psi dx dy - \int_{\mathcal{D}} \mathcal{E}_{\bar{l}} \delta(\phi) \psi dx dy \\ &- \nu \int_{\partial\Omega} \psi \nabla \cdot (\nabla\phi) \delta(\phi) - (\nabla\psi \cdot \nabla\phi) \delta(\phi) + \delta(\phi) \frac{\nabla\phi \cdot \nabla\psi}{\|\nabla\phi\|} dx dy \\ &= - \int_{\mathcal{D}} \mathcal{E}_l \delta(\phi_l) \psi dx dy - \int_{\mathcal{D}} \mathcal{E}_{\bar{l}} \delta(\phi_l) \psi dx dy \\ &+ \nu \int_{\partial\Omega} \nabla \cdot \left(\frac{\nabla\phi_l}{\|\nabla\phi_l\|} \right) \delta(\phi_l) \psi dx dy. \end{aligned} \quad (\text{A.31})$$

Note that $\delta(\phi)$ is not differentiable in the strict sense. We may consider the $\delta(\phi)$ as an approximate Dirac delta in the above derivation. Then, we can apply the Γ -convergence [101] to prove that the derived results hold true when $\delta(\phi)$ is the Dirac delta.

To minimize the functional of (A.30), we do not require performing the curve evolution. Instead, we can update the signed distance function directly.

$$\frac{\partial \phi_l}{\partial t} = (E_{\bar{l}} - E_l) \delta(\phi_l) - \nu \nabla \cdot \left(\frac{\nabla \phi_l}{\|\nabla \phi_l\|} \right) \delta(\phi_l). \quad (\text{A.32})$$

★ End of example ★

An alternative to the calculus of variations for minimizing the energy defined by domain integral is to calculate the shape derivatives. One may refer to [102] for more details.

B Appendix of Chapter 3

Proof of Proposition 3.1

Proof Since $\langle \vec{F}(C^*), \vec{T}(C^*) \rangle = 0$, we can directly obtain

$$\vec{N}(C^*) = \frac{\vec{F}(C^*)}{\|\vec{F}(C^*)\|},$$

or

$$\vec{F}(C^*) = \vec{0},$$

which is one of the definitions of the level set, and this completes the proof.

Proof of Proposition 3.2

Proof Substituting (3.10) and $\frac{dp}{dt}$ into $\frac{dC}{dt}$, we obtain the following:

$$\begin{aligned} \frac{dC}{dt} &= \frac{\partial C}{\partial t} + \frac{\partial C}{\partial p} \frac{dp}{dt} \\ &= \langle \vec{F}, \vec{T} \rangle \vec{N} - \langle \vec{F}, \vec{N} \rangle \vec{T} \\ &= \langle \vec{F}, \mathcal{R}^T \vec{N} \rangle \vec{N} + \langle \vec{F}, \mathcal{R}^T \vec{T} \rangle \vec{T} \\ &= \langle \mathcal{R} \vec{F}, \vec{N} \rangle \vec{N} + \langle \mathcal{R} \vec{F}, \vec{T} \rangle \vec{T} \\ &= -\mathcal{R} \nabla g(p, t), \end{aligned} \tag{B.1}$$

where $\mathcal{R}$ is a 90° rotation matrix of size 2×2 . Thus, taking derivative of g w.r.t. t , we obtain the following:

$$\frac{dg}{dt} = \nabla g(p, t) \frac{dC}{dt} = -\langle \nabla g(p, t), \mathcal{R} \nabla g(p, t) \rangle = 0, \tag{B.2}$$

which completes the proof. ■

Useful elementary results

In the following, some geometrical properties of general curve evolutions are presented, which are used in the analysis of the convergence of curve evolutions.

Theorem B.1 *Given a curve evolution $C(p, t)$ defined by the following:*

$$\begin{aligned}\partial_t C &= \epsilon(p, t) \vec{N}(p, t) \\ \frac{\partial \vec{N}}{\partial t} &= \frac{d\vec{N}}{dt},\end{aligned}\tag{B.3}$$

where the equation for $\vec{N}$ implies the deformation of $\vec{N}$ only depends on the deformation of C , the following holds:

$$\frac{\partial^{(k+1)} C}{\partial t^{(k+1)}} = \frac{\partial^k \epsilon(p, t)}{\partial t^k} \vec{N}(p, t).\tag{B.4}$$

To prove the above theorem, the following lemma is needed.

Lemma B.1

$$\frac{d\vec{N}}{dC}(C^*) \vec{N}(C^*) = 0.\tag{B.5}$$

Proof Firstly, $\frac{d\vec{N}}{dC}(C^*)$ can be written more explicitly as follows:

$$\frac{d\vec{N}}{dC}(C^*) = \begin{pmatrix} \frac{dN^x}{dx}(C^*) & \frac{dN^y}{dx}(C^*) \\ \frac{dN^x}{dy}(C^*) & \frac{dN^y}{dy}(C^*) \end{pmatrix}.$$

According to the relation $\frac{\nabla \Phi}{\|\nabla \Phi\|} = \vec{N}$ and $\|\nabla \Phi\| = 1$ for signed distance function Φ defined by $C^* = \{x, y | \Phi(x, y) = 0\}$, the following is true:

$$\nabla^2 \Phi = \frac{d\vec{N}}{dC}(C^*),$$

i.e., $\frac{d\vec{N}}{dC}(C^*)$ is the Hessian matrix of Φ at the position C^* .

Taking derivative of $\|\nabla\Phi\| = 1$ leads to the following:

$$\nabla\|\nabla\Phi\| = (\nabla^2\Phi)\nabla\Phi = \vec{0},$$

which completes the proof. $\blacksquare$

Now Theorem B.1 can be proven as follows.

Proof of Theorem B.1 The theorem is proven by induction.

For $k = 1$, we have

$$\partial_t C = \epsilon(p, t) \vec{N}(p, t). \quad (\text{B.6})$$

Taking derivative w.r.t t , we obtain the following:

$$\frac{\partial^2 C}{\partial t^2} = \frac{\partial \epsilon(p, t)}{\partial t} \vec{N}(p, t) + \epsilon(p, t) \frac{d\vec{N}}{dC} \frac{\partial C}{\partial t}. \quad (\text{B.7})$$

Recall Lemma B.1 and Eq. (B.6), we obtain the following:

$$\frac{\partial^2 C}{\partial t^2} = \frac{\partial \epsilon(p, t)}{\partial t} \vec{N}(p, t).$$

Now we set $k = n$, we have

$$\frac{\partial^{(n+1)} C}{\partial t^{(n+1)}} = \frac{\partial \epsilon^{(n)}(p, t)}{\partial t^n} \vec{N}(p, t).$$

Hence, taking derivative and by Lemma B.1, we obtain

$$\frac{\partial^{(n+2)} C}{\partial t^{(n+2)}} = \frac{\partial \epsilon^{(n+1)}(p, t)}{\partial t^{(n+1)}} \vec{N}(p, t),$$

which completes the proof. $\blacksquare$.

The above results eventually lead us to a more practical result that is stated as follows.

Theorem B.2 *If there exists $R > 0$, such that $\forall |\tau| < R$, the Taylor series of $C(p, t + \tau)$ at $C(p, t)$ exists, then there exists a unique $\xi(p, \tau)$, such that*

$$C(p, t + \tau) = C(p, t) + \xi(p, \tau) \vec{N}(p, t),$$

where

$$\xi(p, \tau) = \sum_{k=1}^{\infty} \frac{\tau^k}{k!} \frac{\partial^{(k-1)} \epsilon(p, t)}{\partial t^{(k-1)}}.$$

Proof The proof is constructive.

Expand $C(p, t + \tau)$ at $C(p, t)$ by Taylor series as follows:

$$C(p, t + \tau) = \sum_{k=0}^{\infty} \frac{\tau^k}{k!} \frac{\partial^k C(p, t)}{\partial t^k}.$$

By Theorem B.1, we can rewrite the above as follows:

$$C(p, t + \tau) = C(p, t) + \left\{ \sum_{k=1}^{\infty} \frac{\tau^k}{k!} \frac{\partial^{(k-1)} \epsilon(p, t)}{\partial t^{(k-1)}} \right\} \vec{N}(p, t).$$

Since the Taylor expansion exists, hence it is unique. By setting

$$\xi(p, \tau) = \sum_{k=1}^{\infty} \frac{\tau^k}{k!} \frac{\partial^{(k-1)} \epsilon(p, t)}{\partial t^{(k-1)}},$$

we complete the proof. ■

In what follows, an idealized curve evolution equation of Snakes is discussed in which the smoothness terms are dropped, i.e. the curve evolution is driven purely by a gradient vector field. In other words, it is assumed that the converged curve is smooth, meaning that the smoothness terms are very small. The simplified curve evolution equation is as follows:

$$\partial_t \hat{C} = \nabla g(\hat{C}) = \langle \nabla g, \vec{N} \rangle \vec{N} + \langle \nabla g, \vec{T} \rangle \vec{T}. \quad (\text{B.8})$$

The geometry of $\hat{C}(p, t)$ is equivalent to $C(p, t)$ that is defined by the following:

$$\frac{\partial_t C}{\partial t} = \langle \nabla g, \vec{N} \rangle \vec{N}. \quad (\text{B.9})$$

Lemma B.2 *Given that the curve evolution $C(p, t)$ defined by Eq. (B.9) near the converged curve $C(p, t^*)$, i.e. $C(p, t^* - \delta t)$ can be expanded by Taylor series at t^* , within a radius of convergence $|\delta t| < R$, where t^* is the time of convergence, and given that the initial curve $C(p, 0)$ is not on maxima of g (minima of $-g$), then the second order derivative on the direction of the normal of the curve, i.e., $\langle \nabla^2 g \vec{N}, \vec{N} \rangle$, is nonpositive. Hence, there will be no maxima of g (minima of $-g$) on the curve C for any time $t > 0$.*

Proof Let us consider the function $gp(t) = g(C(p, t))$ for any fixed p , then by substituting in Eq. (B.9), $dgp(t)/dt = \|\langle \nabla g, \vec{N} \rangle\|^2 \geq 0$. Hence, g is monotonically nondecreasing w.r.t t . Assuming $C(p, t^*)$ is the converged solution of Eq. (B.9) where $\langle \nabla g, \vec{N} \rangle \Big|_{t=t^*} = 0$, we have that $g(C(p, t^*)) \geq g(C(p, t^* - \delta t))$, $\delta t > 0$.

By Theorem B.2, there exists $\xi(p) = \xi(p, \delta t)$ such that the following relation between $C(p, t^* - \delta t)$ and $C(p, t^*)$ holds:

$$C(p, t^* - \delta t) = C(p, t^*) + \xi(p) \vec{N}(C(p, t^*)).$$

The Taylor expansion of g near the converged solution $C(p, t^* - \delta t) = C(p, t^*) +$

$\xi(p)\vec{N}(C(p, t^*))$, can be written as follows:

$$\begin{aligned}
& g(C(p, t^* - \delta t)) \\
&= g(C(p, t^*)) + \xi \left. \langle \nabla g, \vec{N} \rangle \right|_{t=t^*} + \frac{\xi^2}{2} \left. \langle \nabla^2 g \vec{N}, \vec{N} \rangle \right|_{t=t^*} \\
&\quad + O(|\xi|^2) \Big|_{t=t^*} \\
&= g(C(p, t^*)) + \frac{\xi^2}{2} \left. \langle \nabla^2 g \vec{N}, \vec{N} \rangle \right|_{t=t^*} \\
&\quad + O(|\xi|^2) \Big|_{t=t^*}.
\end{aligned}$$

Since $g(C(p, t^*)) \leq g(C(p, t^* - \delta t))$, we have

$$\left. \langle \nabla^2 g \vec{N}, \vec{N} \rangle \right|_{t=t^*} \geq 0,$$

which completes the proof. ■

Proof of the main results

Proof of Theorem 3.1 The objective of this proof is to obtain a differential equation that defines a contraction mapping. We then can show that the η approaches 0.

According to Taylor's theorem, we write the curve evolution equation near the equilibrium as follows:

$$\begin{aligned}
\left. \frac{\partial C}{\partial t} \right|_{t=t^*} &= \vec{0} \\
&= \left. \frac{\partial C}{\partial t} \right|_{t=t^*-\tau} + \tau \left. \frac{\partial}{\partial t} \frac{\partial C}{\partial t} \right|_{t=t^*-\tau} + O(\tau).
\end{aligned} \tag{B.10}$$

This equation is the basic form for us to construct the contraction mapping. The main idea is to use η to replace the C in this equation to obtain the differential equation of contraction.

We may need the following relationships:

$$t = t^* \Leftrightarrow \tau = 0, \text{ and } \frac{dt}{d\tau} = \frac{d\tau}{dt} = -1.$$

By differentiating Eq. (3.14), we obtain the following:

$$\frac{\partial \eta(p, \tau)}{\partial \tau} = - \left. \frac{\partial C}{\partial t} \right|_{t=t^*-\tau} (-1) = f\vec{N}, \quad (\text{B.11})$$

where $f\vec{N}$ is an dummy term. The form of $f\vec{N}$ is from Lemma 3.1. This relationship allows us to replace C with η for one of the terms in Eq. (B.10).

Now we focus on replacing the C with η for the other term in Eq. (B.10). With the help of Lemma B.1, we have the following derivation:

$$\begin{aligned} \left. \frac{\partial}{\partial t} \frac{\partial C}{\partial t} \right|_{t=t^*-\tau} &= \left. \frac{\partial}{\partial t} \langle \vec{F}(C), \vec{N}(C) \rangle \vec{N}(C) \right|_{t=t^*-\tau} \\ &= \left\langle J[\vec{F}](C) \frac{\partial C}{\partial t}, \vec{N}(C) \right\rangle \vec{N}(C) \Big|_{t=t^*-\tau} \\ &\quad + \left\langle \vec{F}(C), \underbrace{\frac{d\vec{N}}{dC} \frac{\partial C}{\partial t}}_{=0} \right\rangle \vec{N}(C) \Big|_{t=t^*-\tau} \\ &\quad + \langle \vec{F}(C), \vec{N}(C) \rangle \underbrace{\frac{d\vec{N}}{dC} \frac{\partial C}{\partial t}}_{=0} \Big|_{t=t^*-\tau} \\ &= [\vec{N}^T J[\vec{F}] \vec{N}](C) \frac{\partial \eta}{\partial \tau}(C) \Big|_{t=t^*-\tau}. \end{aligned} \quad (\text{B.12})$$

The Taylor expansion of η gives us the following:

$$\begin{aligned} \eta(p, 0) &= \eta(p, \tau) - \tau \frac{\partial \eta}{\partial \tau} + O(\tau) \\ \Leftrightarrow \eta(p, \tau) &= \tau \frac{\partial \eta}{\partial \tau} + O(\tau), \end{aligned}$$

where $\eta(p, 0) = 0$.

Substituting the above into Eq.(B.12), we obtain the following:

$$\begin{aligned} \frac{\partial C}{\partial t} \Big|_{t=t^*-\tau} &= -\tau \frac{\partial}{\partial t} \frac{\partial C}{\partial t} \Big|_{t=t^*-\tau} + O(\tau) \\ &= -[\vec{N}^T J[\vec{F}] \vec{N}](C) \eta(p, \tau) + O(\tau). \end{aligned}$$

Combining the above with Eq. (B.11), we finally arrive at the following:

$$\begin{aligned} \frac{\partial \eta}{\partial \tau} \Big|_{\tau=t^*-t} &= -\left[\vec{N}^T(C) J[\vec{F}](C) \vec{N}(C) \right] \eta(p, \tau) \\ &\quad + O(\tau) \\ &= -J_{nn}(C) \eta(p, \tau) + O(\tau), \end{aligned} \tag{B.13}$$

where

$$J_{nn}(C) = [\vec{N}^T J[\vec{F}] \vec{N}](C). \tag{B.14}$$

Now, we almost obtain the contraction of η except that the curve $C = C(p, t^* - \tau)$ is not convenient to deal with. Therefore, the following approximation is used:

$$\begin{aligned} J_{nn}(C) &= [\vec{N}^T J[\vec{F}] \vec{N}](C^*) - \tau \left[\vec{N}^T \left[\frac{dJ[\vec{F}]}{dC} \otimes \frac{\partial C}{\partial t} \right] \vec{N} \right] + O(\tau) \\ &\approx [\vec{N}^T J[\vec{F}] \vec{N}](C^*), \end{aligned} \tag{B.15}$$

where $\frac{dJ[\vec{F}]}{dC}$ is the second order derivative of the vector field, which is relatively small comparing to $J[\vec{F}]$. This treatment is valid for the case where the first order derivative does not vanish. Otherwise, it is categorized as the *marginally stable*.

To further explain how Eq.(B.13) with the approximate in Eq. (B.15) can define a contraction mapping, and thus can be used for test of convergence, we discretize the differential equation and neglect the higher order terms $O(\tau)$ to obtain the

following:

$$\begin{aligned}\eta^k(p) &= \eta^{k-1}(p) -_{\Delta} \tau J_{nn}(C^*) \eta^{k-1}(p) \\ &= (1 -_{\Delta} \tau J_{nn}(C^*)) \eta^{k-1}(p).\end{aligned}$$

Then we take norm of the above to obtain the following:

$$\begin{aligned}\|\eta^k(p)\| &= |1 -_{\Delta} \tau J_{nn}(C^*)| \|\eta^{k-1}(p)\| \\ &= |1 -_{\Delta} \tau J_{nn}(C^*)|^k \|\eta^0(p)\|.\end{aligned}$$

Noting that $\tau = t^* - t$ is decreasing, then $_{\Delta}\tau < 0$. Hence, i.i.f $|1 +_{\Delta} \tau J_{nn}(C^*)| < 1$, $\lim_{t \rightarrow \infty} \|\eta^k(p)\| = 0$, i.e., η will vanish. And i.i.f $J_{nn}(C^*) < 0$, $|1 +_{\Delta} \tau J_{nn}(C^*)| < 1$. ■

Proof of Theorem 3.4 The vector field $\vec{F}$ is defined below:

$$\vec{F} = (A^T + \epsilon I) \nabla g.$$

The Jacobian matrix is therefore the following:

$$\begin{aligned}\vec{N}^T J[\vec{F}] \vec{N} &= \vec{N}^T (A^T + \epsilon I) J[\nabla g] \vec{N} \\ &= \vec{N}^T (A^T J[\nabla g]) \vec{N} + \vec{N}^T (\epsilon J[\nabla g]) \vec{N} \\ &= 0 + \epsilon g_{nn} \leq 0.\end{aligned}$$

The above together with Theorem 3.1 completes the proof. ■

C Appendix of Chapter 4

Proof of Proposition 4.4 Since $H_l \in \{0, 1\}$ and $\sum_{l=1}^2 H_l = 1$, we have $\sum_{l=1}^2 \mathcal{E}_l(x, y)H_l = \{\mathcal{E}_1(x, y), \mathcal{E}_2(x, y)\}$. Thus, for any position $[x, y]^T$ in Ω , we can choose $\mathcal{E}_{l^*}$ from $\{\mathcal{E}_1(x, y), \mathcal{E}_2(x, y)\}$ by $\mathcal{E}_{l^*} = \min\{\mathcal{E}_1(x, y), \mathcal{E}_2(x, y)\}$, which is equivalent to

$$\{H_l\}^* = \underset{\{H_l\}}{\operatorname{argmin}} \left\{ \sum_{l=1}^2 \mathcal{E}_l(x, y)H_l \right\}$$

at $[x, y]^T$. As a result, $\sum_{l=1}^2 \mathcal{E}_l(x, y)H_l^* \leq \sum_{l=1}^2 \mathcal{E}_l(x, y)H_l$ for any H_l . Thus,

$$\int_{\Omega} \sum_{l=1}^2 \mathcal{E}_l(x, y)H_l^* dx dy \leq \int_{\Omega} \sum_{l=1}^2 \mathcal{E}_l(x, y)H_l dx dy,$$

which completes this proof. $\blacksquare$

Note that the proof relies on the Axiom of Choice for the continuous domain. The choice is feasible if we approximate the integral by discretization.

Proof of Proposition 4.5 To prove this proposition, we require the following lemma.

Lemma C.1 *Given the same condition in Proposition 4.5 then the following holds:*

$$\begin{aligned} \varepsilon_{seg} \leq & \frac{\int_{\Omega} \mathcal{E}_1(x, y)H dx dy}{|\Omega_1|\mathcal{E}_2(x', y')} \\ & + \frac{\int_{\Omega} \mathcal{E}_2(x, y)(1 - H) dx dy}{|\Omega_2|\mathcal{E}_1(x', y')}. \end{aligned} \tag{C.1}$$

The proof of this lemma is due to Markov's inequality. This bound connects the error of segmentation to the error of reconstruction.

Additionally, we require the following fundamental model of functions in the literature of signal analysis, such as wavelets [103] and especially compressive sensing [104, 105].

Definition 7.2 (Universal energy bound) *Given a (discrete) function $f \in \mathbb{R}^N$, and any subset of fixed system of sorted orthogonal bases $\{v_i | i = 1, 2, \dots, K \leq N\}$ such that $\|\langle f, v_1 \rangle\|_{l_2} \geq \|\langle f, v_2 \rangle\|_{l_2} \geq \dots \geq \|\langle f, v_K \rangle\|_{l_2}$, then the following bound holds for every $0 < n \leq N$:*

$$\|\langle f, v_n \rangle\|_{l_2} \leq Rn^{-q}, \quad (\text{C.2})$$

where $R = \|\langle f, v_1 \rangle\|_{l_2}$, and $0 \leq q < \infty$.

The worst case is when $\|\langle f, v_1 \rangle\|_{l_2} = \|\langle f, v_2 \rangle\|_{l_2} = \dots = \|\langle f, v_K \rangle\|_{l_2}$, i.e. $q = 0$.

From the definition of universal energy bound, the linear patch reconstruction error at every pixel could be bounded by using (C.2) as follows:

$$\begin{aligned} \|\boxtimes \|\mathcal{E}_l^\boxtimes &= \left\| \mathbf{p} - \sum_{k=1}^K \langle \mathbf{p}, \mathbf{v}_k \rangle_{\boxtimes} \mathbf{v}_k \right\|_{\boxtimes}^2 \\ &= \|\mathbf{p}\|_{\boxtimes}^2 - \sum_{k=1}^K \langle \mathbf{p}, \mathbf{v}_k \rangle_{\boxtimes}^2 \\ &\geq \|\mathbf{p}\|_{\boxtimes}^2 - \sum_{k=1}^K R_l^2 k^{-2q_l} \\ &\geq \|\mathbf{p}_l\|_{\boxtimes}^2 - R_l^2 K^{2-2q_l}, \end{aligned} \quad (\text{C.3})$$

where $l = 1, 2$. Substituting (C.3) into the denominator of (C.1), we complete our proof. $\blacksquare$

Theorem C.1 *The optimization problem defined in Eq. (4.11) is nonconvex.*

Proof of Theorem C.1 First, we shall prove that the objective functional is non-

convex. Taking functional derivatives of the energy twice, we have the following:

$$\begin{aligned}
& \frac{D^2}{D\mathbf{v}_i^2} E_p(\{\mathbf{v}_i\}) \\
&= - \int_{\Omega} I(x+u', y+v') I(x+u, y+v) H_l dx dy \\
&= - \int_{\Omega} I(x+u', y+v') H_l I(x+u, y+v) H_l dx dy \quad (\text{C.4}) \\
&= - \int_{\Omega} I_{Hl}(x+u', y+v') I_{Hl}(x+u, y+v) dx dy \\
&= -\Lambda_{Hl}(u', v', u, v),
\end{aligned}$$

where $I_{Hl} = IH_l$. If this functional is convex, $-\Lambda_{Hl}(u', v', u, v)$ will be positive semi-definite. However, we can show that the $-\Lambda_{Hl}(u', v', u, v)$ is actually negative semi-definite as follows:

$$\begin{aligned}
& - \int_{\Omega} h(u', v') \Lambda_{Hl}(u', v', u, v) h(u, v) du dv du' dv' \\
&= - \int_{\Omega} \left| \int_{\boxtimes} I_{Hl}(x+u, y+v) h(u, v) du dv \right|^2 dx dy \quad (\text{C.5}) \\
&\leq 0.
\end{aligned}$$

Hence, the objective functional is non-convex but concave. Further, we consider the constraints in terms of orthonormality. Let $\{\mathbf{x}_i\}$ and $\{\mathbf{y}_i\}$ both satisfy the constraints. We wonder whether the convex combination also satisfy the constraints. For example, let $\mathbf{w}_1 = \alpha \mathbf{x}_1 + (1-\alpha) \mathbf{y}_1$, where $\alpha > 0$. Then we can verify that $\|\mathbf{w}_1\| \neq 1$, unless $\mathbf{x}_1 = \mathbf{y}_1$, which is generally not true. Therefore, the constraints are also nonconvex. Either of the above two facts can complete our proof. ■

Theorem C.2 *In the optimization problem defined in Eq. (4.11), if $\{\mathbf{v}_k, k = 1, 2, \dots, K\}$ is the global optimal solution, then the transformation of $\{\mathbf{v}_k, k =$*

$1, 2, \dots, K\}$ via an orthogonal matrix $\mathcal{R}^{K \times K}$ is also the global optimal solution.

Proof of Theorem C.2 First, let us define a linear transformed patch basis $\{\mathbf{w}_k\}$ as follows:

$$\begin{aligned}\mathbf{w}_1(u, v) &= [\mathbf{v}_1^l(u, v), \mathbf{v}_2^l(u, v), \dots, \mathbf{v}_K^l(u, v)]^T \mathbf{a}_1 \\ \mathbf{w}_2(u, v) &= [\mathbf{v}_1^l(u, v), \mathbf{v}_2^l(u, v), \dots, \mathbf{v}_K^l(u, v)]^T \mathbf{a}_2 \\ &\vdots \\ \mathbf{w}_K(u, v) &= [\mathbf{v}_1^l(u, v), \mathbf{v}_2^l(u, v), \dots, \mathbf{v}_K^l(u, v)]^T \mathbf{a}_K,\end{aligned}\tag{C.6}$$

where $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_K]^T$ is an $K \times K$ mixing matrix. We may simply write the following:

$$\mathbf{w}_1 = \mathbf{V}\mathbf{a}_1, \mathbf{w}_2 = \mathbf{V}\mathbf{a}_2, \dots, \mathbf{w}_K = \mathbf{V}\mathbf{a}_K,\tag{C.7}$$

where $\mathbf{V} = [\mathbf{v}_1^l(u, v), \mathbf{v}_2^l(u, v), \dots, \mathbf{v}_K^l(u, v)]^T$.

Hence, the energy in terms of $\mathbf{w}_k(u, v)$ could be rewritten as follows:

$$\begin{aligned}U(\{\mathbf{w}_k\}) &= \int_{\Omega} \sum_{k=1}^K \langle \mathbf{p}, \mathbf{w}_k \rangle_{\boxtimes}^2 H_l dx dy \\ &= \int_{\Omega} \sum_{k=1}^K \langle \mathbf{p}, \mathbf{V}\mathbf{a}_k \rangle_{\boxtimes}^2 H_l dx dy \\ &= \int_{\Omega} \|\mathbf{A}\mathbf{V}^T \mathbf{p}\|^2 H_l dx dy \\ &= \int_{\Omega} (\mathbf{p}^T \mathbf{V} [\mathbf{A}^T \mathbf{A}] \mathbf{V}^T \mathbf{p}) H_l dx dy.\end{aligned}\tag{C.8}$$

If the mixing matrix $\mathbf{A}$ is orthogonal, the following is true:

$$U(\{\mathbf{w}_i\}) = \int_{\Omega} (\mathbf{p}^T \mathbf{V} \mathbf{V}^T \mathbf{p}) H_l dx dy = U(\{\mathbf{v}_i^l\}).\tag{C.9}$$

Therefore, if $\{\mathbf{v}_i^l\}$ is the global optimal solution, then $\{\mathbf{w}_i\}$ is also the global optimal

solution, which completes the proof. $\blacksquare$

To prove Theorem 4.7 we require some lemmas.

Lemma C.2 *The integral operator $[\Lambda_{HI}]$ is symmetric positive semi-definite.*

The symmetry is straightforward. The proof of positive semi-definiteness is as in the proof of Theorem C.1.

Hence according to Mercer's theorem of eigen-decomposition of symmetric non-negative definite bounded integral operator, we can write $\Lambda_{HI}(u, v, u', v')$ in the form of infinite series of eigenfunctions as follows.

Theorem C.3 (Mercer's representation) *Assuming $\|[\Lambda_{HI}(u, v, u', v')]\| < \infty$, then*

$$\Lambda_{HI}(u, v, u', v') = \sum_{i=1}^{\infty} \lambda_i \mathbf{e}_i(u, v) \mathbf{e}_i(u', v'), \quad (\text{C.10})$$

where $\lambda_1 \geq \lambda_2 \geq \dots > 0$ are the eigenvalues, $\{\mathbf{e}_i(u, v), i = 1, 2, \dots\}$ is the set of eigenfunctions, and the eigenfunctions form a system of orthogonal basis.

A proof of the above can be found in [106]. Now, we are in a position to prove the Theorem 4.7.

Proof of Theorem 4.7: First, we write the gradient descent differential equation for the STEP 1 as follows:

$$\begin{aligned} & \frac{\partial \mathbf{v}_1^l(u, v, t)}{\partial t} \\ &= \int \Lambda_{HI}(u, v, u', v') \mathbf{v}_1^l(u', v', t) du' dv', \end{aligned} \quad (\text{C.11})$$

where $\mathbf{v}_1^l(u, v, 0) = v_0$. The corresponding update equation is the following:

$$\begin{aligned} \mathbf{v}_1^{l,n+1}(u, v) &= \mathbf{v}_1^{l,n}(u, v) \\ &+ \Delta_t \int \Lambda_{HI} \mathbf{v}_1^{l,n}(u', v') du' dv'. \end{aligned} \quad (\text{C.12})$$

Suppose $V_0 = \sum_{i=1}^{\infty} \alpha_i \mathbf{e}_i(u, v)$, the update equation could be rewritten as follows:

$$\begin{aligned}
 \mathbf{v}_1^{l,1}(u, v) &= \sum_{i=1}^{\infty} \alpha_i \mathbf{e}_i(u, v) \\
 &\quad + \Delta_t \int \Lambda_{Hl} \sum_{i=1}^{\infty} \alpha_i \mathbf{e}_i(u, v) du' dv' \\
 &= \sum_{i=1}^{\infty} \alpha_i \mathbf{e}_i(u, v) + \Delta_t \sum_{i=1}^{\infty} \lambda_i \alpha_i \mathbf{e}_i(u, v) \\
 &= \sum_{i=1}^{\infty} (1 + \Delta_t \lambda_i) \alpha_i \mathbf{e}_i(u, v),
 \end{aligned} \tag{C.13}$$

where we applied Mercer's representation. Hence,

$$\begin{aligned}
 \mathbf{v}_1^{l,n}(u, v) &= \sum_{i=1}^{\infty} (1 + \Delta_t \lambda_i)^n \alpha_i \mathbf{e}_i(u, v) \\
 &= (1 + \Delta_t \lambda_1)^n \left[\alpha_1 \mathbf{e}_1(u, v) \right. \\
 &\quad \left. + \sum_{i=2}^{\infty} \underbrace{\left(\frac{1 + \Delta_t \lambda_i}{1 + \Delta_t \lambda_1} \right)^n}_{\approx 0} \alpha_i \mathbf{e}_i(u, v) \right] \\
 &\approx (1 + \Delta_t \lambda_1)^n \alpha_1 \mathbf{e}_1(u, v).
 \end{aligned} \tag{C.14}$$

The normalization constraint gives us the desired result:

$$\lim_{n \rightarrow \infty} \mathbf{v}_1^{l,n}(u, v) = \frac{\mathbf{v}_1^{l,n}(u, v)}{\|\mathbf{v}_1^l(u, v)\|} = \pm \mathbf{e}_1(u, v), \tag{C.15}$$

where we may omit the $\pm$ sign. The proofs for the subsequent steps is hence straightforward, where we only need to replace $\mathbf{v}_0 = \sum_{i=1}^{\infty} \alpha_i \mathbf{e}_i(u, v)$ with $\mathbf{v}_0 = \sum_{i=j}^{\infty} \alpha_i \mathbf{e}_i(u, v)$, $j = 2, 3, \dots, K$. This setting is due to the Gram-Schmidt process in the constraints. The energy $U(\{\mathbf{v}_i^l\})$ is therefore the sum of eigenvalues due to Mercer's representation, which completes this proof. ■

Proof of Corollary 4.2 According to the proof of Theorem 4.7, the iteration for

STEP k is the following:

$$\begin{aligned} \mathbf{v}_k^n(u, v) &= (1 + \Delta_t \lambda_k)^n \left[\alpha_k \mathbf{e}_k(u, v) \right. \\ &\quad \left. + \sum_{h=k+1}^N \left(\frac{1 + \Delta_t \lambda_h}{1 + \Delta_t \lambda_k} \right)^n \alpha_h \mathbf{e}_h(u, v) \right]. \end{aligned} \quad (\text{C.16})$$

Let $\tilde{\mathbf{v}}_k^n(u, v) = \mathbf{v}_k^n(u, v) / (1 + \Delta_t \lambda_k)^n$, $\beta_h = \left\| \frac{1 + \Delta_t \lambda_h}{1 + \Delta_t \lambda_k} \right\|$. Note that the eigenvalues are ordered. Hence, $\beta_h < 1$. From these we deduce the following:

$$\begin{aligned} \frac{\|\tilde{\mathbf{v}}_k^{n+1}(u, v) - \alpha_k \mathbf{e}_k(u, v)\|}{\|\tilde{\mathbf{v}}_k^n(u, v) - \alpha_k \mathbf{e}_k(u, v)\|} &= \frac{\sum_{h=k+1}^N \beta_h^{n+1} |\alpha_h|}{\sum_{h=k+1}^N \beta_h^n |\alpha_h|} \\ &= 1 - \frac{\sum_{h=k+1}^N \beta_h^n |\alpha_h| (1 - \beta_h)}{\sum_{h=k+1}^N \beta_h^n |\alpha_h|} \\ &\leq 1 - \sum_{h=k+1}^N \frac{\beta_h^n |\alpha_h| (1 - \beta_h)}{(N - k - 1) \beta_*^n |\alpha_*|} \\ &\leq 1 - \frac{1}{N - k - 1} \sum_{h=k+1}^N (1 - \beta_h) \\ &= \frac{1}{N - k - 1} \sum_{h=k+1}^N \beta_h < 1, \end{aligned} \quad (\text{C.17})$$

where $\beta_*^n |\alpha_*| = \max\{\beta_h^n |\alpha_h|, h = k + 1, i + 2, \dots, N\}$. The above completes the proof. $\blacksquare$

Proof of Theorem 4.6: The energy of Eq. (4.12) in terms of $\{\mathbf{w}_i, i = 1, 2, \dots, k\}$ is the following:

$$\begin{aligned} U(\{\mathbf{w}_k\}) &= \int_{\Omega} \sum_{k=1}^K \langle \mathbf{p}, \mathbf{w}_k \rangle_{\boxtimes}^2 dx dy \\ &= \int_{\{u, v, u', v'\}} \sum_{k=1}^K \mathbf{w}_k(u, v) \mathbf{w}_k(u', v') \Lambda_{Hl} du dv du' dv' \\ &= \sum_{k=1}^K \sum_{h=1}^N \alpha_{kh}^2 \lambda_h, \end{aligned} \quad (\text{C.18})$$

where $\mathbf{w}_k = \sum_{h=1}^N \alpha_{kh} \mathbf{e}_h$. Note that $\sum_{h=1}^N \alpha_{kh}^2 = 1$. In words, the summation over h is a convex combination of λ_h . Note that the λ_h are ordered from large values to small values for $h = 1$ to N . Hence, the following holds:

$$\begin{aligned}
 \sum_{h=1}^N \alpha_{kh}^2 \lambda_h &= \sum_{h=1}^K \alpha_{kh}^2 \lambda_h + \sum_{h'=K+1}^N \alpha_{kh'}^2 \lambda_{h'} \\
 &= \sum_{h=1}^K \alpha_{kh}^2 \lambda_h + \sum_{h'=K+1}^N \alpha_{kh'}^2 \left(\sum_{h=1}^K \beta_h \right) \lambda_{h'} \\
 &\quad \text{‘where, } \sum_{h=1}^K \beta_h = 1, \text{ and, } \beta_h > 0 \\
 &\leq \sum_{h=1}^K \alpha_{kh}^2 \lambda_h + \sum_{h'=K+1}^N \alpha_{kh'}^2 \left(\sum_{h=1}^K \beta_h \lambda_h \right).
 \end{aligned} \tag{C.19}$$

The last inequality is due to the fact that the convex combination of larger values is larger than smaller values. Now, we consider a special case of β_h as follows:

$$\beta_h = \frac{1 - \sum_{k=1}^K \alpha_{kh}^2}{\sum_{k=1}^K \left(1 - \sum_{h=1}^K \alpha_{kh}^2 \right)}. \tag{C.20}$$

This β_h satisfies the following:

$$\sum_{h=1}^K \beta_h = \frac{\sum_{h=1}^K \left(1 - \sum_{k=1}^K \alpha_{kh}^2 \right)}{\sum_{k=1}^K \left(1 - \sum_{h=1}^K \alpha_{kh}^2 \right)} = 1. \tag{C.21}$$

Besides, regarding the inequality in (C.19) we have the following:

$$\begin{aligned}
& \sum_{k=1}^K \left(\alpha_{kh}^2 + \beta_h \sum_{h'=K+1}^N \alpha_{kh'}^2 \right) \\
&= \sum_{k=1}^K \left[\alpha_{kh}^2 + \beta_h \left(1 - \sum_{h'=K}^N \alpha_{kh'}^2 \right) \right] \\
&= \sum_{k=1}^K \left[\alpha_{kh}^2 + \frac{1 - \sum_{k=1}^K \alpha_{kh}^2}{\sum_{k=1}^K \left(1 - \sum_{h=1}^K \alpha_{kh}^2 \right)} \left(1 - \sum_{h'=1}^K \alpha_{kh'}^2 \right) \right] \quad (\text{C.22}) \\
&= \sum_{k=1}^K \left[\alpha_{kh}^2 + \frac{1 - \sum_{h'=1}^K \alpha_{kh'}^2}{\sum_{k=1}^K \left(1 - \sum_{h=1}^K \alpha_{kh}^2 \right)} \left(1 - \sum_{k=1}^K \alpha_{kh}^2 \right) \right] \\
&= \sum_{k=1}^K \alpha_{kh}^2 + \left(1 - \sum_{k=1}^K \alpha_{kh}^2 \right) = 1.
\end{aligned}$$

Substituting the above into (C.19), we have the following:

$$\begin{aligned}
U(\{\mathbf{w}_k\}) &\leq \sum_{h=1}^K \lambda_h \sum_{k=1}^K \left(\alpha_{kh}^2 + \beta_h \sum_{h'=K+1}^N \alpha_{kh'}^2 \right) \\
&= \sum_{h=1}^K \lambda_h = U(\{\mathbf{e}_k\}), \quad (\text{C.23})
\end{aligned}$$

which completes our proof.

The global optimality of PCA for the reconstruction of zero-mean vectors has been reported in [107]. The proof of global optimality of the eigenpatches here for reconstruction of arbitrary patches is quite different from theirs, and the procedure of the proof here is simpler.

D Appendix of Chapter 5

Proof of Theorem 5.8 Let us now consider the PVSE discretized in time as follows:

$$C(s, t + \Delta t) = C(s, t) + \Delta t \sum_i^n \frac{d\theta_i}{dt} \mathbf{V}^{\theta_i}(C(s, t)), \quad (\text{D.1})$$

where we parameterize the curve by arclength for simplicity. Taking derivatives with respect to s twice we may obtain Eq. (D.2) for the curvature, where $C_{ss} = \kappa \mathbf{N}$, and κ is the curvature.

$$\begin{aligned} C_{ss}(s, t + \Delta t) &= C_{ss}(s, t) + \Delta t \sum_i^n \frac{d\theta_i}{dt} \frac{\partial}{\partial s} \left(\left[\frac{\mathbf{D}\mathbf{V}^{\theta_i}}{\mathbf{D}\mathbf{x}} \right]^{2 \times 2} C_s \right) \\ &= C_{ss}(s, t) + \Delta t \sum_i^n \frac{d\theta_i}{dt} \left\{ C_s \otimes \left[\frac{\mathbf{D}^2 \mathbf{V}^{\theta_i}}{\mathbf{D}C^2} \right]^{2 \times 2 \times 2} \otimes C_s + \left[\frac{\mathbf{D}\mathbf{V}^{\theta_i}}{\mathbf{D}\mathbf{x}} \right] C_{ss} \right\}. \end{aligned} \quad (\text{D.2})$$

Let us consider the position $C(s, t)$, such that $\kappa(s, t) = 0$. We wish to know how κ changes at these feature points. From Eq. (D.2), we obtain the following:

$$\begin{aligned} C_{ss}(s, t + \Delta t) &= C_{ss}(s, t) + \Delta t \sum_i^n \frac{d\theta_i}{dt} \left\{ \mathbf{T} \otimes \left[\frac{\mathbf{D}^2 \mathbf{V}^{\theta_i}}{\mathbf{D}\mathbf{x}^2} \right] \otimes \mathbf{T} \right\}, \end{aligned} \quad (\text{D.3})$$

where $\mathbf{T} = C_s$ is the tangent vector of the contour.

Further rearrangement yields the expression for $\frac{\partial C_{ss}}{\partial t}$ as follows:

$$\frac{\partial C_{ss}}{\partial t} = \sum_i^n \frac{d\theta_i}{dt} \left\{ \mathbf{T} \otimes \left[\frac{\mathbf{D}^2 \mathbf{V}^{\theta_i}}{\mathbf{D}\mathbf{x}^2} \right] \otimes \mathbf{T} \right\}. \quad (\text{D.4})$$

Moreover, since

$$\frac{\partial C_{ss}}{\partial t} = \frac{\partial \kappa \mathbf{N}}{\partial t} = \kappa_t \mathbf{N} + \underbrace{\kappa}_{=0} \mathbf{N}_t = \kappa_t \mathbf{N}, \quad (\text{D.5})$$

we have $\left\| \frac{\partial C_{ss}}{\partial t} \right\| = \|\kappa_t\|$, which leads to the following:

$$\begin{aligned} \left\| \frac{\partial \kappa}{\partial t} \right\| &= \left\| \sum_i^n \frac{d\theta_i}{dt} \left\{ \mathbf{T} \otimes \left[\frac{\mathbf{D}^2 \mathbf{V}^{\theta_i}}{\mathbf{D} \mathbf{x}^2} \right] \otimes \mathbf{T} \right\} \right\| \\ &\leq \left\| \frac{d\theta}{dt} \right\| \sum_i^n \left\| \frac{\mathbf{D}^2 \mathbf{V}^{\theta_i}}{\mathbf{D} \mathbf{x}^2} \right\|, \end{aligned} \quad (\text{D.6})$$

and this completes the proof. $\blacksquare$

Proof of Corollary 5.4 We begin with the following identity:

$$\left. \frac{dp}{dt} \right|_{\kappa=0} = \left. \frac{dp}{ds} \frac{ds}{d\kappa} \frac{\partial \kappa}{\partial t} \right|_{\kappa=0} = \frac{1}{\|C_p\| \kappa_s} \left. \frac{\partial \kappa}{\partial t} \right|_{\kappa=0}, \quad (\text{D.7})$$

where we applied $ds = \|C_p\| dp$. Note that $\left. \frac{dp}{dt} \right|_{\kappa(p,t)=0}$ is different from the $\frac{dp}{dt}$ at a fixed p . Rather, $\left. \frac{dp}{dt} \right|_{\kappa(p,t)=0}$ defines on a moving p . Besides, $\frac{ds}{d\kappa}$ is possible for nonzero κ_s , since $\kappa : s \mapsto \kappa(s)$ is bijection.

Hence, we can establish the following bound:

$$\left| \left. \frac{dp}{dt} \right|_{\kappa=0} \right| \leq \frac{1}{\|\kappa_s C_p\|} \left\| \frac{d\theta}{dt} \right\| \sum_i^n \left\| \frac{\mathbf{D}^2 \mathbf{V}^{\theta_i}}{\mathbf{D} \mathbf{x}^2} \right\|. \quad (\text{D.8})$$

For the two feature points $C(p_1, t)$ and $C(p_2, t)$ at time t , we have the following:

$$\begin{aligned} \left| \left. \frac{d}{dt} |p_1 - p_2| \right| \right| &\leq 2 |\text{sign}(p_1 - p_2)| \left| \left. \frac{dp}{dt} \right|_{\kappa=0} \right| \\ &\leq \frac{2}{\|\kappa_s C_p\|} \left\| \frac{d\theta}{dt} \right\| \sum_i^n \left\| \frac{\mathbf{D}^2 \mathbf{V}^{\theta_i}}{\mathbf{D} \mathbf{x}^2} \right\|, \end{aligned} \quad (\text{D.9})$$

which completes the proof. $\blacksquare$

Proof of Theorem 5.9 To induce the norm of $\frac{\mathbf{D}^2 \mathbf{V}^{\theta}}{\mathbf{D} \mathbf{x}^2}$, we expand the derivative

as follows:

$$\begin{aligned} & \left[\frac{\mathbf{D}^2 e_{mn}^1}{\mathbf{D}\mathbf{x}^2}, \frac{\mathbf{D}^2 e_{mn}^2}{\mathbf{D}\mathbf{x}^2} \right] \\ &= \left\{ \begin{bmatrix} \frac{\partial^2 e_{mn}^1}{\partial x^2} & \frac{\partial^2 e_{mn}^1}{\partial y \partial x} \\ \frac{\partial^2 e_{mn}^1}{\partial x \partial y} & \frac{\partial^2 e_{mn}^1}{\partial y^2} \end{bmatrix}, \right. \\ & \quad \left. \begin{bmatrix} \frac{\partial^2 e_{mn}^2}{\partial x^2} & \frac{\partial^2 e_{mn}^2}{\partial y \partial x} \\ \frac{\partial^2 e_{mn}^2}{\partial x \partial y} & \frac{\partial^2 e_{mn}^2}{\partial y^2} \end{bmatrix} \right\}^{[2 \times 2 \times 2]}, \end{aligned} \quad (\text{D.10})$$

where $1 \leq m \leq M, 1 \leq n \leq N$. Therefore, we may bound the norm as follows:

$$\begin{aligned} & \left\| \left[\frac{\mathbf{D}^2 e_{mn}^1}{\mathbf{D}\mathbf{x}^2}, \frac{\mathbf{D}^2 e_{mn}^2}{\mathbf{D}\mathbf{x}^2} \right] \right\| \leq \left\| \frac{\mathbf{D}^2 e_{mn}^1}{\mathbf{D}\mathbf{x}^2} \right\| + \left\| \frac{\mathbf{D}^2 e_{mn}^2}{\mathbf{D}\mathbf{x}^2} \right\| \\ &= \left\| \begin{bmatrix} \frac{\partial^2 e_{mn}^1}{\partial x^2} & \frac{\partial^2 e_{mn}^1}{\partial y \partial x} \\ \frac{\partial^2 e_{mn}^1}{\partial x \partial y} & \frac{\partial^2 e_{mn}^1}{\partial y^2} \end{bmatrix} \right\| + \left\| \begin{bmatrix} \frac{\partial^2 e_{mn}^2}{\partial x^2} & \frac{\partial^2 e_{mn}^2}{\partial y \partial x} \\ \frac{\partial^2 e_{mn}^2}{\partial x \partial y} & \frac{\partial^2 e_{mn}^2}{\partial y^2} \end{bmatrix} \right\|. \end{aligned} \quad (\text{D.11})$$

Hence, we may consider to bound $\left\| \frac{\mathbf{D}^2 e_{mn}^1}{\mathbf{D}\mathbf{x}^2} \right\|_2, \left\| \frac{\mathbf{D}^2 e_{mn}^2}{\mathbf{D}\mathbf{x}^2} \right\|$ separately. For $\left\| \frac{\mathbf{D}^2 e_{mn}^1}{\mathbf{D}\mathbf{x}^2} \right\|$, if we choose the $\|\cdot\| = \|\cdot\|_\infty$, we have the following bound:

$$\begin{aligned} & \left\| \frac{\mathbf{D}^2 e_{mn}^1}{\mathbf{D}\mathbf{x}^2} \right\| = \left\| \begin{bmatrix} \frac{\partial^2 e_{mn}^1}{\partial x^2} & \frac{\partial^2 e_{mn}^1}{\partial y \partial x} \\ \frac{\partial^2 e_{mn}^1}{\partial x \partial y} & \frac{\partial^2 e_{mn}^1}{\partial y^2} \end{bmatrix} \right\| \\ &= \max \left(\left| \frac{\partial^2 e_{mn}^1}{\partial x^2} + \frac{\partial^2 e_{mn}^1}{\partial y \partial x} \right|, \left| \frac{\partial^2 e_{mn}^1}{\partial x \partial y} + \frac{\partial^2 e_{mn}^1}{\partial y^2} \right| \right) \\ &\leq \max \left(\frac{n^2 + mn}{n^2 + m^2}, \frac{m^2 + mn}{n^2 + m^2} \right) \\ &\leq \frac{2}{1+c} \leq 2, \end{aligned} \quad (\text{D.12})$$

where $c = m/n$ if $m \leq n$ and $c = n/m$ if $n < m$. Hence,

$$\sum_{mn} \left\| \frac{\mathbf{D}^2 e_{mn}^i}{\mathbf{D}\mathbf{x}^2} \right\| = \sum_{m=1}^M \sum_{n=1}^N \left\| \frac{\mathbf{D}^2 e_{mn}^i}{\mathbf{D}\mathbf{x}^2} \right\| \leq 2MN. \quad (\text{D.13})$$

---

for $i = 1, 2$. The above together with (D.11) completes the proof. ■

E Appendix of Chapter 6

Gradient descent equations for shape training

The gradient descent equations for learning the parameters $\{\alpha_i\}$ and β in the shape decision function in Eq. (6.4) are as follows:

$$\begin{aligned}\frac{\partial \alpha_i}{\partial t} &= - \int_{\mathcal{D}} 2(H_e - H_o) \psi(\|\vec{z} - \vec{p}_i\|) H' dx dy, \\ i &= 1, 2, \dots, N,\end{aligned}$$

$$\frac{\partial \beta}{\partial t} = - \int_{\mathcal{D}} 2(H_e - H_o) H' dx dy,$$

where H' is the first order derivative of H . The trained contour curve is defined by Eq. (6.2) in which H_o is replaced by the trained H_e .

Proof of Proposition 6.8

Proof Let $\vec{a} = \vec{z}_c - \vec{p}_i$, we have that $\|\mathbf{A}\vec{a}\|^2 = \vec{a}^T \mathbf{A}^T \mathbf{A} \vec{a}$. Since $\mathbf{A}^T \mathbf{A} \neq \mathbf{I}$ by non-orthogonality, we have $\|\mathbf{A}\vec{a}\|^2 - \|\vec{a}\|^2 = \vec{a}^T (\mathbf{A} - \mathbf{I})^T (\mathbf{A} - \mathbf{I}) \vec{a}$, and $\mathbf{A} - \mathbf{I}$ is a nonzero matrix. Hence, $\|\mathbf{A}\vec{a}\|^2 - \|\vec{a}\|^2 \neq 0$ if $(\mathbf{A} - \mathbf{I})\vec{a} \neq 0$. The latter inequality has infinitely many solutions, which completes the proof. ■

Derivations of Eq. (6.12)

In the following, we present the derivations of Eq. (6.12).

$$\begin{aligned}\frac{dE}{dt} &= - \nabla_{\mathbf{A}} E \left(\nabla_{\mathbf{A}} J - \left\langle \nabla_{\mathbf{A}} J, \frac{\nabla_{\mathbf{A}} E}{\|\nabla_{\mathbf{A}} E\|} \right\rangle \frac{\nabla_{\mathbf{A}} E}{\|\nabla_{\mathbf{A}} E\|} \right) \\ &\quad - \nabla_{\vec{b}} E \left(\nabla_{\vec{b}} J - \left\langle \nabla_{\vec{b}} J, \frac{\nabla_{\vec{b}} E}{\|\nabla_{\vec{b}} E\|} \right\rangle \frac{\nabla_{\vec{b}} E}{\|\nabla_{\vec{b}} E\|} \right) \\ &= - \langle \nabla_{\mathbf{A}} J, \nabla_{\mathbf{A}} E \rangle + \langle \nabla_{\mathbf{A}} J, \nabla_{\mathbf{A}} E \rangle \\ &\quad - \langle \nabla_{\vec{b}} J, \nabla_{\vec{b}} E \rangle + \langle \nabla_{\vec{b}} J, \nabla_{\vec{b}} E \rangle = 0.\end{aligned}$$

$$\begin{aligned}
\frac{dJ}{dt} &= -\nabla_{\mathbf{A}} J \left(\nabla_{\mathbf{A}} J - \left\langle \nabla_{\mathbf{A}} J, \frac{\nabla_{\mathbf{A}} E}{\|\nabla_{\mathbf{A}} E\|} \right\rangle \frac{\nabla_{\mathbf{A}} E}{\|\nabla_{\mathbf{A}} E\|} \right) \\
&\quad - \nabla_{\vec{b}} J \left(\nabla_{\vec{b}} J - \left\langle \nabla_{\vec{b}} J, \frac{\nabla_{\vec{b}} E}{\|\nabla_{\vec{b}} E\|} \right\rangle \frac{\nabla_{\vec{b}} E}{\|\nabla_{\vec{b}} E\|} \right) \\
&= - \left(\|\nabla_{\mathbf{A}} J\|^2 - \left\| \left\langle \nabla_{\mathbf{A}} J, \frac{\nabla_{\vec{b}} E}{\|\nabla_{\vec{b}} E\|} \right\rangle \right\|^2 \right) \\
&\quad - \left(\|\nabla_{\vec{b}} J\|^2 - \left\| \left\langle \nabla_{\vec{b}} J, \frac{\nabla_{\vec{b}} E}{\|\nabla_{\vec{b}} E\|} \right\rangle \right\|^2 \right) \leq 0.
\end{aligned}$$

Derivation of Eqs. (6.13) and (6.14)

Our derivation is based on the following equality for minimizing general active contour energy:

$$\frac{\partial C}{\partial t} = -\nabla J = \alpha(p)\vec{N}. \quad (\text{E.1})$$

Note that the equality holds true for geometric active contours.

The differential of the shape decision function ϕ_S at the implicit contour leads to the following:

$$\begin{aligned}
&\left. \frac{\partial \phi_S(\vec{z}(t), \{\mathbf{A}^{-1}(t), \vec{b}(t)\})}{\partial t} \right|_{\vec{z} \in \{\vec{z} | \phi_S = 0\}} \\
&= \nabla_{\vec{z}} \phi_S \cdot \frac{\partial C}{\partial t} + \frac{D\phi_S}{D\mathbf{A}^{-1}} \cdot \frac{d\mathbf{A}^{-1}}{dt} + \frac{D\phi_S}{D\vec{b}} \cdot \frac{d\vec{b}}{dt} = 0.
\end{aligned} \quad (\text{E.2})$$

Substituting Eq. (E.1) into the above, we obtain the expression for α :

$$\alpha = - \frac{\frac{D\phi_S}{D\mathbf{A}^{-1}} \cdot \frac{d\mathbf{A}^{-1}}{dt} + \frac{D\phi_S}{D\vec{b}} \cdot \frac{d\vec{b}}{dt}}{\langle \vec{N}, \nabla_{\vec{z}} \phi_S \rangle}. \quad (\text{E.3})$$

Substituting (E.3) into (E.1) we obtain the curve evolution as follows:

$$\frac{\partial C}{\partial t} = - \frac{\frac{D\phi_S}{D\mathbf{A}^{-1}} \cdot \frac{d\mathbf{A}^{-1}}{dt} + \frac{D\phi_S}{D\vec{b}} \cdot \frac{d\vec{b}}{dt}}{\langle \vec{N}, \nabla_{\vec{z}} \phi_S \rangle} \vec{N}. \quad (\text{E.4})$$

To minimize a general active contour energy $J(C)$, we require the derivative of $J(C)$ to be non-positive as follows:

$$\frac{dJ}{dt} = \left\langle \nabla J, \frac{\partial C}{\partial t} \right\rangle_p \leq 0, \quad (\text{E.5})$$

where $\langle f, g \rangle_p = \int \langle f(p), g(p) \rangle dp$ is the inner product of two vector functions in which $\langle, \rangle$ is the vector inner product. Substituting (E.3) into (E.5), and by considering $\vec{N} = \nabla_{\vec{z}} \phi_S$, we obtain the following:

$$\frac{dJ}{dt} = \left\langle \nabla J, -\frac{\frac{D\phi_S}{D\mathbf{A}^{-1}} \cdot \frac{d\mathbf{A}^{-1}}{dt} + \frac{D\phi_S}{D\vec{b}} \cdot \frac{d\vec{b}}{dt} \vec{N}}{\langle \vec{N}, \nabla_{\vec{z}} \phi_S \rangle} \vec{N} \right\rangle_p \leq 0, \quad (\text{E.6})$$

where $\frac{d\mathbf{A}^{-1}}{dt} = -\nabla_{\mathbf{A}^{-1}} J$ and $\frac{d\vec{b}}{dt} = -\nabla_{\vec{b}} J$. We may verify that Eqs. (6.13) and (6.14) ensure that Eq. (E.6) is valid.

Author's Publication

- (a) Junyan Wang, Kap Luk Chan, and Y. Wang, On the stationary solution of PDE based curve evolution, in Proceedings of the 19th British Machine Vision Conference, 2008.
- (b) Junyan Wang and Kap Luk Chan, Shape evolution for rigid and nonrigid shape registration and recovery, in IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2009.

Bibliography

- [1] D. Marr, *Vision*. W. H. Freeman and Company, 1980.
- [2] T. Poggio, V. Torre, and C. Koch, “Computational vision and regularization theory,” *Nature*, vol. 317, pp. 314–319, 1985.
- [3] M. Kass, A. Witkin, and D. Terzopoulos, “Snakes: Active contour models,” *International Journal of Computer Vision*, vol. 1, no. 4, pp. 321–331, 1988.
- [4] D. Marr and E. Hildreth, “Theory of edge detection,” in *Proceedings of the Royal Society of London. Series B, Biological Sciences*, vol. 207, no. 1167, 1980, pp. 187–217.
- [5] C. Xu and J. L. Prince, “Snakes, shapes, and gradient vector flow,” *IEEE Transactions on Image Processing*, vol. 7, no. 3, pp. 359–369, 1998.
- [6] A. J. Yezzi, S. Kichenassamy, A. Kumar, P. J. Olver, and A. Tannenbaum, “A geometric snake model for segmentation of medical imagery,” *IEEE Transactions on Medical Imaging*, vol. 16, no. 2, pp. 199–209, 1997.
- [7] W. Fang, K. L. Chan, S. Fu, and S. Krishnan, “Incorporating temporal information into level set functional for robust ventricular boundary detection from echocardiographic image sequence,” *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 11, pp. 2548–2556, 2008.

- [8] N. Paragios and R. Deriche, “Geodesic active contours and level sets for the detection and tracking of moving objects,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 3, pp. 266–280, 2000.
- [9] R. Goldenberg, R. Kimmel, E. Rivlin, and M. Rudzsky, “Fast geodesic active contours,” in *Proceedings of the Second International Conference on Scale-Space Theories in Computer Vision*. London, UK: Springer-Verlag, 1999, pp. 34–45.
- [10] W. Fang and K. L. Chan, “Incorporating shape prior into geodesic active contours for detecting partially occluded object,” *Pattern Recognition*, vol. 40, no. 8, pp. 2163–2172, 2007.
- [11] G. Sapiro, *Geometric Partial Differential Equations and Image Analysis*. New York, NY, USA: Cambridge University Press, 2006.
- [12] S. Osher and N. Paragios, *Geometric Level Set Methods in Imaging, Vision, and Graphics*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2003.
- [13] S. Osher and R. Fedkiw., *Level-Set Methods and Dynamic Implicit Surfaces*. Springer, New York, 2003.
- [14] S. Osher and J. A. Sethian, “Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations,” *Journal of Computational Physics*, vol. 79, pp. 12–49, 1988.
- [15] V. Caselles, F. Catté, T. Coll1, and F. Dibos, “A geometric model for active contours in image processing,” *Numerische Mathematik*, vol. 66(1), pp. 1–31, 1993.

- [16] R. Malladi, J. A. Sethian, and B. C. Vemuri, "Shape modeling with front propagation: A level set approach," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, no. 2, pp. 158–175, 1995.
- [17] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contour," *International Journal of Computer Vision*, vol. 22, no. 1, pp. 61–79, 1997.
- [18] S. C. Zhu and A. Yuille, "Region competition: Unifying snakes, region growing, and Bayes/MDL for multiband image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, no. 9, pp. 884–900, 1996.
- [19] R. Ronfard, "Region-based strategies for active contour models," *International Journal of Computer Vision*, vol. 13, no. 2, pp. 229–251, 1994.
- [20] D. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and associated variational problems," *Communications on Pure and Applied Mathematics*, vol. 42, no. 5, pp. 577–685, 1989.
- [21] A. Blake and A. Zisserman, *Visual reconstruction*, 1987.
- [22] T. Pavlidis, "Segmentation of pictures and maps through functional approximation," *Computer Graphics and Image Processing*, vol. 1, pp. 360–372, 1972.
- [23] S. Zucker, "Region growing: Childhood and adolescence," *Computer graphics and image processing*, vol. 5, no. 3, pp. 382–399, 1976.
- [24] T. Chan and L. Vese, "Active contours without edges," *IEEE Transactions on Image Processing*, vol. 10, no. 2, pp. 266–277, 2001.

- [25] L. Grady and C. V. Alvino, “The piecewise smooth mumford-shah functional on an arbitrary graph,” *IEEE Transactions on Image Processing*, vol. 18, no. 11, pp. 2547–2561, 2009.
- [26] E. S. Brown, T. F. Chan, and X. Bresson, “Globally convex chan-ese image segmentation,” UCLA CAM Report, Tech. Rep., 2010.
- [27] T. Pock, D. Cremers, H. Bischof, and A. Chambolle, “An algorithm for minimizing the mumford-shah functional,” in *Proceedings of the 12th IEEE International Conference on Computer Vision*, 2009.
- [28] T. Brox and D. Cremers, “On local region models and a statistical interpretation of the piecewise smooth mumford-shah functional,” *International Journal of Computer Vision*, vol. 84, no. 2, pp. 184–193, 2009.
- [29] B. Sumengen and B. Manjunath, “Graph partitioning active contours (gpac) for image segmentation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, pp. 509–521, 2006.
- [30] J. Yuan, E. Bae, and X. Tai, “A study on continuous max-flow and min-cut approaches,” in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2010.
- [31] Y. Chen, H. D. Tagare, S. Thiruvenkadam, F. Huang, D. Wilson, K. S. Gopinath, R. W. Briggs, and E. A. Geiser, “Using prior shapes in geometric active contours in a variational framework,” *International Journal of Computer Vision*, vol. 50, no. 3, pp. 315–328, 2002.
- [32] M. Leventon, W. Grimson, and O. Faugeras, “Statistical shape influence in geodesic active contours,” in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2000.

- [33] D. Cremers, T. Kohlberger, and C. Schnörr, “Shape statistics in kernel space for variational image segmentation,” *Pattern Recognition*, vol. 36, no. 9, pp. 1929–1943, 2003.
- [34] D. Cremers, S. J. Osher, and S. Soatto, “Kernel density estimation and intrinsic alignment for shape priors in level set segmentation,” *International Journal of Computer Vision*, vol. 3175, pp. 36–44, 2004.
- [35] P. Etyngier, F. Segonne, and R. Keriven, “Shape priors using manifold learning techniques,” in *Proceedings of the Eleventh IEEE International Conference on Computer Vision*, 2007.
- [36] T. Schoenemann and D. Cremers, “A combinatorial solution for model-based image segmentation and real-time tracking,” *TPAMI*, vol. 32, pp. 1153–1164, 2010.
- [37] D. Mumford and J. Shah, “Boundary detection by minimizing functionals,” in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 1985.
- [38] L. C. Evans and R. F. Gariepy, *Measure Theory and Fine Properties of Functions*. CRC Press, 1992.
- [39] C. Sagiv, N. Sochen, and Y. Zeevi, “Integrated active contours for texture segmentation,” *IEEE Transactions on Image Processing*, vol. 16, no. 6, pp. 1633–1646, 2006.
- [40] S. Lankton and A. Tannenbaum, “Localizing Region-Based Active Contours,” *IEEE Transactions on Image Processing*, vol. 17, no. 11, pp. 2029–2039, 2008.

- [41] C. Li, C. Kao, J. C. Gore, and Z. Ding, “Minimization of region-scalable fitting energy for image segmentation,” *IEEE Transactions on Image Processing*, vol. 17, no. 10, pp. 1940–1949, 2008.
- [42] N. Paragios, O. Mellina-Gottardo, and V. Ramesh, “Gradient vector flow fast geometric active contours,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 3, pp. 402–407, 2004.
- [43] X. Xie and M. Mirmehdi, “Mac: Magnetostatic active contour model,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 4, pp. 632–646, 2008.
- [44] L. D. Cohen, “On active contour models and balloons,” *CVGIP: Image Understanding*, vol. 53, no. 2, pp. 211–218, 1991.
- [45] C. Li, J. Liu, and M. D. Fox, “Segmentation of edge preserving gradient vector flow: An approach toward automatically initializing and splitting of snakes,” in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2005.
- [46] N. Paragios and R. Deriche, “Geodesic active regions and level set methods for supervised texture segmentation,” *The International Journal of Computer Vision*, vol. 46, no. 3, pp. 223–247, 2002.
- [47] C. Xu, A. Yezzi, and J. L. Prince, “On the relationship between parametric and geometric active contours,” in *Asilomar Conference on Signals, Systems, and Computers*, vol. 1, Oct. 2000, pp. 483–489.
- [48] C. L. Epstein and M. Gage, “The curve shortening flow,” in *Wave Motion: Theory, Modeling, and Computation*, A. Chorin and A. Majda, Eds. Springer-Verlag, New York, 1987.

-
- [49] H. Khalil, *Nonlinear Systems*. Prentice Hall, 2002.
- [50] N. Paragios, O. Mellina-Gottardo, and V. Ramesh, “Gradient vector flow fast geodesic active contours,” in *Proceedings of the 11th IEEE International Conference on Computer Vision*, 2001.
- [51] C. Li, C. Xu, C. Gui, and M. D. Fox, “Distance regularized level set evolution and its application to image segmentation,” *IEEE Transactions on Image Processing*, vol. 19, no. 12, pp. 3243–3254, December 2010.
- [52] D. Adalsteinsson and J. A. Sethian, “A fast level set method for propagating interfaces,” *Journal of Computational Physics*, vol. 118, no. 2, pp. 269–277, 1995.
- [53] J. A. Sethian, *Level Set Methods and Fast Marching Methods: Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision, and Materials Science*. Cambridge University Press, June 1999.
- [54] Y. Shi and W. Karl, “A real-time algorithm for the approximation of level-set-based curve evolution,” *IEEE Transactions on Image Processing*, vol. 17, no. 5, pp. 645–656, 2008.
- [55] B. Sandberg, T. Chan, and L. Vese, “A level-set and Gabor-based active contour algorithm for segmenting textured images,” *UCLA CAM Report*, 2002.
- [56] I. Kokkinos, G. Evangelopoulos, and P. Maragos, “Texture analysis and segmentation using modulation features, generative models, and weighted curve evolution,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 1, pp. 142–157, 2009.

- [57] M. Rousson, T. Brox, and R. Deriche, “Active unsupervised texture segmentation on a diffusion based feature space,” in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2003.
- [58] N. Houhou, J.-P. Thiran, and X. Bresson, “Fast Texture Segmentation Model based on the Shape Operator and Active Contour,” in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2008.
- [59] T. S. Lee, D. Mumford, and A. L. Yuille, “Texture segmentation by minimizing vector-valued energy functionals: The coupled-membrane model,” in *Proceedings of the Second European Conference on Computer Vision*, 1992.
- [60] L. A. Vese and T. F. Chan, “A multiphase level set framework for image segmentation using the mumford and shah model,” *International Journal of Computer Vision*, vol. 50, no. 3, pp. 271–293, 2002.
- [61] J. Piovano, M. Rousson, and T. Papadopoulos, “Efficient segmentation of piecewise smooth images,” in *Proceedings of the 1st international conference on Scale space and variational methods in computer vision*, 2007.
- [62] J. B. Rosen, “Global minimization of a linearly constrained concave function by partition of feasible domain,” *Mathematics of Operations Research*, vol. 8, no. 2, pp. 215–230, May 1983.
- [63] A. Tsai, A. J. Y. Jr., and A. S. Willsky, “Curve evolution implementation of the mumford-shah functional for image segmentation, denoising, interpolation, and magnification,” *IEEE Transactions on Image Processing*, vol. 10, no. 8, pp. 1169–1186, 2001.
- [64] J. C. Bezdek and R. J. Hathaway, “Convergence of alternating optimization,” *Neural Parallel and Scientific Computations*, vol. 11, no. 4, pp. 351–368, Dec. 2003.

- [65] T. Randen and J. H. Husøy, “Filtering for texture classification: A comparative study,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, pp. 291–310, 1999.
- [66] P. Brodatz, *Textures: A Photographic Album for Artists and Designers*. New York, NY, Dover, 1996, <http://www.ux.uis.no/~tranden/brodatz.html>.
- [67] K. Ni, X. Bresson, T. Chan, and S. Esedoglu, “Local histogram based segmentation using the wasserstein distance,” *International Journal of Computer Vision*, vol. 84, no. 1, pp. 97–111, Aug. 2009.
- [68] S. Roth and M. J. Black, “Fields of experts,” *International Journal of Computer Vision*, vol. 82, no. 2, pp. 205–229, 2009.
- [69] D. Martin, C. Fowlkes, D. Tal, and J. Malik, “A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics,” in *Proceedings of the 8th International Conference on Computer Vision*, vol. 2, July 2001, pp. 416–423.
- [70] Y. Boykov and V. Kolmogorov, “Computing geodesics and minimal surfaces via graph cuts,” in *Proceedings of the 9th IEEE International Conference on Computer Vision*, 2003.
- [71] M. Rousson and N. Paragios, “Shape priors for level set representations,” in *Proceedings of the 7th European Conference on Computer Vision*, 2002.
- [72] S. H. Joshi and A. Srivastava, “Intrinsic bayesian active contours for extraction of object boundaries in images,” *International Journal of Computer Vision*, vol. 81, no. 3, pp. 331–355, 2009.

- [73] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models-their training and application," *Computer Vision and Image Understanding*, vol. 61, no. 1, pp. 38–59, Jan. 1995.
- [74] Y. Amit, U. Grenander, and M. Piccioni, "Structural image restoration through deformable templates," *Journal of the American Statistical Association*, vol. 86, no. 414, pp. 376–387, 1991.
- [75] A. Jain, Y. Zhong, and S. Lakshmanan, "Object matching using deformable templates," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, no. 3, pp. 267–278, 1996.
- [76] C. A. Glasbey and K. V. Mardia, "A review of image-warping methods," *Journal of Applied Statistics*, vol. 25, no. 2, pp. 155–171, 1998.
- [77] P. F. Felzenszwalb, "Representation and detection of deformable shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 2, pp. 208–220, 2005.
- [78] Y. N. Wu, Z. Si, C. Fleming, and S.-C. Zhu, "Deformable template as active basis," in *Proceedings of 11th IEEE International Conference on Computer Vision*, 2007.
- [79] A. Mansouri, D. P. Mukherjee, and S. T. Acton, "Constraining active contour evolution via lie groups of transformation," *IEEE Transactions on Image Processing*, vol. 13, pp. 853–863, 2004.
- [80] D. P. Mukherjee and S. T. Acton, "Affine and projective active contour models," *Pattern Recognition*, vol. 40, no. 3, pp. 920–930, 2007.

-
- [81] G. Mori, S. Belongie, and J. Malik, “Efficient shape matching using shape contexts,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 11, pp. 1832–1837, November 2005.
  - [82] S. Abbasi, F. Mokhtarian, and J. Kittler, “Curvature scale space image in shape similarity retrieval,” *Multimedia Systems*, vol. 7, no. 6, pp. 467–476, 1999.
  - [83] F. Mokhtarian and M. Bober, *Curvature Scale Space Representation: Theory, Applications, and MPEG-7 Standardization*. Norwell, MA, USA: Kluwer Academic Publishers, 2003.
  - [84] L. Staib and J. Duncan, “Boundary finding with parametrically deformable models,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, pp. 1061–1075, 1992.
  - [85] S. Krinidis and I. Pitas, “Fast free-vibration modal analysis of 2-D physics-based deformable objects,” *IEEE Transactions on Image Processing*, vol. 14, no. 3, pp. 281–293, 2005.
  - [86] J. Wang and K. L. Chan, “Shape evolution for rigid and nonrigid shape registration and recovery,” in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2009.
  - [87] C. Xu and J. L. Prince, “Snakes, shapes, and gradient vector flow,” *IEEE Transactions on Image Processing*, vol. 7, no. 3, pp. 359–369, 1998.
  - [88] B. Li and S. T. Acton, “Automatic active model initialization via poisson inverse gradient,” *IEEE Transactions on Image Processing*, vol. 17, no. 8, pp. 1406–1420, 2008.

- [89] D. Cremers, F. R. Schmidt, and F. Barthel, “Shape priors in variational image segmentation: Convexity, lipschitz continuity and globally optimal solutions,” in *IEEE Conference on Computer Vision and Pattern Recognition*, 2008.
- [90] T. Schoenemann and D. Cremers, “A combinatorial solution for model-based image segmentation and real-time tracking,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, pp. 1153–1164, 2010.
- [91] P. Viola and M. J. Jones, “Robust real-time face detection,” *International Journal of Computer Vision*, vol. 57, pp. 137–154, May 2004.
- [92] N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection,” in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2005.
- [93] D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.
- [94] H. Jiang and S. Yu, “Linear solution to scale and rotation invariant object matching,” in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2009.
- [95] H. Li, E. Kim, X. Huang, and L. He, “Object matching with a locally affine-invariant constraint.” in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2010.
- [96] R. Kimmel, K. Siddiqi, B. B. Kimia, and A. Bruckstein, “Shape from shading: Level set propagation and viscosity solutions,” *International Journal of Computer Vision*, vol. 16, no. 2, pp. 107–133, 1995.

- [97] A. Kumar, A. R. Tannenbaum, and G. J. Balas, “Optical flow: A curve evolution approach,” *IEEE Transactions on Image Processing*, vol. 5, pp. 598–610, 1996.
- [98] P. J. Olver, “Applied Mathematics Lecture Notes.” [Online]. Available: <http://www.math.umn.edu/~olver/appl.html>
- [99] R. Weinstock, *Calculus of Variations, with Applications to Physics and Engineering*. New York: Dover, 1974.
- [100] P. Y. Papalambros and D. J. Wilde, *Principles of Optimal Design*. New York, NY, USA: Cambridge University Press, 2003.
- [101] A. Braides, *Γ -convergence for beginners*. Oxford University Press, 2002.
- [102] S. Jehan-Besson, M. Barlaud, and G. Aubert, “Dream2s: Deformable regions driven by an eulerian accurate minimization method for image and video segmentation,” *International Journal of Computer Vision*, vol. 53, no. 1, pp. 45–70, 2003.
- [103] Y. Meyer, *Wavelets and operators*, ser. Cambridge Studies in Advanced Mathematics. Cambridge: Cambridge University Press, 1992, translated from the 1990 French original by D. H. Salinger.
- [104] D. L. Donoho, “Compressed sensing,” *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
- [105] E. J. Candes and T. Tao, “Near-optimal signal recovery from random projections: Universal encoding strategies?” *IEEE Transactions on Information Theory*, vol. 52, no. 12, pp. 5406–5425, 2006.
- [106] R. B. Ash, *Information Theory*. Wiley, New York, 1965.

-
- [107] P. Baldi and K. Hornik, “Neural networks and principal component analysis: Learning from examples without local minima,” *Neural Networks*, vol. 2, no. 1, pp. 53 – 58, 1989.